Semantically Secure Lattice Codes for the Gaussian Wiretap Channel

Cong Ling, Laura Luzzi, Jean-Claude Belfiore, and Damien Stehlé

***Abstract*—We prove that nested lattice codes can achieve semantic security and strong secrecy over the Gaussian wiretap channel. The key tool in our proof is the flatness factor which characterizes the convergence of the conditional output distributions corresponding to different messages and leads to an upper bound on the information leakage. We not only show the existence of lattice codes that are good for secrecy, but also propose the flatness factor as a new design criterion. Both the modulo-lattice Gaussian channel and the genuine Gaussian channel are considered. In the latter case, we propose a new secrecy coding scheme based on the discrete Gaussian distribution over a lattice, which achieves the secrecy capacity to within a half nat under mild conditions. No *a priori* distribution of the message is assumed, and no dither is used in our proposed schemes.**

***Index Terms*—lattice coding, physical layer security, strong secrecy, semantic security, wiretap channel.**

I. Introduction

The idea of information-theoretic security stems from Shannon's notion of *perfect secrecy*. Perfect security can be achieved by encoding an information message M (also called plaintext message), belonging to a finite space $\mathcal{M}$, into a codeword or ciphertext Z, belonging to a discrete or continuous space $\mathcal{Z}$, in such a way that the mutual information $\mathbb{I}(\mathsf{M};\mathsf{Z}) = 0$. However, perfect security is impractical because it requires a one-time pad.

In the context of noisy channels, Wyner [1] proved that both robustness to transmission errors and a prescribed degree of data confidentiality could simultaneously be attained by channel coding without any secret key. Wyner replaced Shannon's perfect secrecy with the *weak secrecy* condition $\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathsf{M};\mathsf{Z}^n) = 0$, namely the asymptotic rate of leaked information between the message M and the channel output Z^n should vanish as the block length n tends to infinity.

Unfortunately, it is still possible for a scheme satisfying weak secrecy to exhibit some security flaws, e.g., the total amount of leaked information may go to infinity, and now it is widely accepted that a physical-layer security scheme should be secure in the sense of Csiszár's *strong secrecy* $\lim_{n\to\infty} \mathbb{I}(\mathsf{M};\mathsf{Z}^n) = 0$ [2].

In the notion of strong secrecy, plaintext messages are often assumed to be random and uniformly distributed in $\mathcal{M}$. This assumption is deemed problematic from the cryptographic perspective, since in many setups plaintext messages are not random. This issue can be resolved by using the standard notion of *semantic security* [3] which requires that the probability that the eavesdropper can guess any function of the message given the ciphertext should not be significantly higher than the probability of guessing it using a simulator that does not have access to the ciphertext. The relation between strong secrecy and semantic security was revealed in [4] for discrete wiretap channels, namely, achieving strong secrecy for all distributions of the plaintext messages is equivalent to achieving semantic security.

Explicit wiretap codes achieving strong secrecy over discrete memoryless channels have been proposed in [5, 6]. In particular, polar codes in [6] also achieve semantic security (although this was implicit in [6]). For continuous channels such as the Gaussian channel, the problem of achieving strong secrecy has been little explored so far and the design of wiretap codes has mostly focused on the maximization of the eavesdropper's error probability [7]. Recently, some progress has been made in using nested lattice codes over Gaussian wiretap channels [8, 9]. It is quite natural to replace Wyner's random binning with coset coding induced by a lattice partition $\Lambda_e \subset \Lambda_b$. The secret bits are used to select one coset of the coarse lattice Λ_e and a random point inside this coset is transmitted. Explicit wiretap lattice codes from an error probability point of view were proposed in [10], which also introduced the notion of *secrecy gain* and showed that the eavesdropper's error probability $\lim_{n\to\infty} P_e = 1$ for even unimodular lattices. These lattice codes were further investigated in [11]. Finally, in [12] the existence of lattice codes (based on the ensemble of random lattice codes) achieving the secrecy capacity under the weak secrecy criterion was demonstrated.

Main Contributions

In the present work, we prove that lattice codes can achieve strong secrecy and semantic security over (continuous) Gaussian wiretap channels.

Firstly, we follow Csiszár's idea [2] to show that strong secrecy is guaranteed if the conditional output distributions

This work was supported in part by FP7 PHYLAWS, by a Royal Society-CNRS international joint project and by a Marie Curie Fellowship (FP7/2007-2013, grant agreement PIEF-GA-2010-274765). This work was presented in part at the IEEE International Symposium on Information Theory (ISIT 2012), Cambridge, MA, USA.

C. Ling is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, United Kingdom (e-mail: cling@ieee.org).

L. Luzzi was with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, United Kingdom. She is now with Laboratoire ETIS (ENSEA - Université de Cergy-Pontoise - CNRS), 6 Avenue du Ponceau, 95014 Cergy-Pontoise, France (e-mail: laura.luzzi@ensea.fr).

Jean-Claude Belfiore is with the Department of Communications and Electronics, Telecom ParisTech, Paris, France (e-mail: belfiore@telecom-paristech.fr).

D. Stehlé is with ENS de Lyon, Laboratoire LIP (U. Lyon, CNRS, ENS de Lyon, INRIA, UCBL), 46 Allée d'Italie, 69364 Lyon Cedex 07, France (e-mail: damien.stehle@ens-lyon.fr).

corresponding to different messages converge to the same distribution in the sense of L^1 distance (also sometimes referred to as *variational distance* or *statistical distance*). This allows us to extend the relation between strong secrecy and semantic security [4] to continuous wiretap channels. More precisely, we derive a bound on the mutual information in terms of the variational distance for continuous channels.

More importantly, we propose the *flatness factor* of a lattice as a fundamental criterion which guarantees L^1 convergence of conditional outputs and characterizes the amount of information leakage. This leads to defining a notion of lattices that are "good for secrecy", similarly to the notions of good lattices which have been proposed for coding problems. Following the approach of Loeliger [13], we then show the existence of infinite families of secrecy-good lattices which can be obtained by lifting linear codes over finite fields using Construction A. To establish this proof, we introduce a modified version of Loeliger's Minkowski-Hlawka theorem, which allows to average a non-compactly supported function over a lattice ensemble.

Before tackling the problem of coding for the Gaussian wiretap channel, and to gain useful insights, we consider a simplified scenario, the *mod-Λ wiretap channel*, and show the existence of nested lattices which guarantee strong secrecy against eavesdroppers and reliability for the legitimate receiver at the same time. The analysis of the mod-Λ channel was a key element in the proof that lattice coding and decoding achieve the capacity of the additive white Gaussian noise (AWGN) channel in [14, 15]. The mod-Λ wiretap channel was already considered in [8] in the context of secrecy with noisy feedback, where it was suggested that finding good wiretap codes for this model can give significant insight to solve the AWGN case. However, as observed in [8] and [9], transferring these techniques from the modulo-Λ case to the AWGN case is not trivial since the modulo structure helps to conceal information from the eavesdropper. We solve this difficulty by employing *lattice Gaussian signaling* for the Gaussian wiretap scenario. More precisely, the distribution of each bin in our wiretap code is a discrete Gaussian distribution over a coset of a secrecy-good lattice. Non-uniform signaling for AWGN channels using discrete Gaussian inputs was already used in [16], where it was shown that such inputs are optimal in terms of shaping gains. Our contribution is to use the flatness factor to show that discrete Gaussian signaling over good lattices can approach the secrecy capacity of the Gaussian wiretap channel up to a constant gap of $\frac{1}{2}$ nat (under very mild assumptions) by using a minimum mean-square error (MMSE) filter at the legitimate receiver.

The proposed approach shows a couple of salient features. Firstly, throughout the paper, we do not make any assumption on the distribution of the plaintext message M, i.e., the security holds for any particular message. Thus, similarly to [4], we prove that lattice codes can achieve semantic security. Secondly, in contrast to what is nowadays the common practice of lattice coding [15], we do not use a dither. This may simplify the implementation of the system.

Relations to Existing Works

Relation to secrecy gain: Given the fundamental volume of a lattice, a small flatness factor requires a small theta series, which coincides with the criterion from [10] for enjoying a large secrecy gain. Thus, although different criteria are adopted in [10] and in this paper, they are in fact consistent with each other.

Relation to resolvability: In [17, 18], a technique based on *resolvability* was suggested to obtain strong secrecy, which uses a binning scheme such that the bin rate is above the capacity of the eavesdropper's channel. We will show this is also the case for the proposed lattice scheme.

Relation to lattice-based cryptography: Lattice-based cryptography [19] aims at realizing classical cryptographic primitives, such as digital signatures and public-key encryption schemes, that are provably secure *under algorithmic hardness assumptions* on worst-case lattice problems, such as variants of the decisional shortest vector problem. In the present work, we propose an encryption scheme for the Gaussian wiretap channel that involves lattices, but the security is proven *without algorithmic hardness assumptions*.

Organization

Section II studies the relation between semantic security and strong secrecy for continuous wiretap channels. In Section III, we review lattice Gaussian distributions and propose the flatness factor. Sections IV and V address the mod-Λ channel and the Gaussian wiretap channel, respectively. In Section VI, we conclude the paper with a brief discussion of open issues.

Throughout this paper, we use the natural logarithm, denoted by $\log$, and information is measured in nats. We use the standard asymptotic notation $f(x) = O(g(x))$ when $\limsup_{x\to\infty} |f(x)/g(x)| < \infty$, $f(x) = \Omega(g(x))$ when $\limsup_{x\to\infty} |g(x)/f(x)| < \infty$, and $f(x) = o(g(x))$ when $\limsup_{x\to\infty} |f(x)/g(x)| = 0$.

II. Strong secrecy and semantic security in continuous channels

In this section, we investigate the relation between strong secrecy and semantic security in continuous wiretap channels. In particular, we extend the results of Section 3 of [4] to such channels. The results are general and apply to more than lattice codes.

A. Wiretap Codes

In this section we briefly recall some basic definitions for the wiretap setting. For a general introduction to continuous wiretap channels, we refer the reader to [20, 21].

Consider an n-dimensional continuous memoryless wiretap channel with input X^n and outputs Y^n, Z^n defined by the i.i.d. conditional distributions:

$$p_{\mathsf{Y}^n|\mathsf{X}^n}(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p_{\mathsf{X}|\mathsf{Y}}(y_i|x_i),$$

$$p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathbf{z}|\mathbf{x}) = \prod_{i=1}^{n} p_{\mathsf{Z}|\mathsf{X}}(z_i|x_i)$$

for the legitimate receiver and the eavesdropper respectively. The random variables $\mathsf{X}^n, \mathsf{Y}^n, \mathsf{Z}^n$ take values in $\mathbb{R}^n$. The channel input is subject to an average power constraint P:

$$\frac{1}{n}\mathbb{E}\left[\|\mathsf{X}^n\|^2\right] \leq P. \tag{1}$$

Definition 1 (Wiretap code)**.** *An (R, R', n) wiretap code for the channel defined above is given by a message set $\mathcal{M}_n = \{1, \ldots, e^{nR}\}$, an auxiliary discrete random source* S *of entropy rate R' taking values in $\mathcal{S}_n$, an encoding function $f_n : \mathcal{M}_n \times \mathcal{S}_n \to \mathbb{R}^n$ and a decoding function $g_n : \mathbb{R}^n \to \mathcal{M}_n$ for the legitimate receiver. Let $\mathsf{X}^n = f_n(\mathsf{M}, \mathsf{S})$ be the channel input for a distribution* M *of messages, and $\hat{\mathsf{M}} = g_n(\mathsf{Y}^n)$ the estimate of the legitimate receiver. The channel input X^n must satisfy the average power constraint (1), with respect to* M *chosen as the uniform distribution and to the randomness source* S*. Alternatively, one can impose a more stringent average power constraint on each individual bin (without assuming* M *is uniformly distributed):*

$$\forall m \in \mathcal{M}_n, \quad \frac{1}{n}\mathbb{E}_{\mathsf{S}}\left[\|f_n(m, \mathsf{S})\|^2\right] \leq P. \tag{2}$$

We denote by $\mathcal{C}_n = f_n(\mathcal{M}_n, \mathcal{S}_n)$ the set of codewords and by $\mathcal{C}_n(m) = f_n(m, \mathcal{S}_n)$ the “bin” corresponding to the message $m \in \mathcal{M}_n$.

Incidentally, the proposed lattice codes will satisfy the individual power constraint.

Remark 1. Note that we do not impose the randomness set $\mathcal{S}_n$ to be finite. In fact, in the scheme of Section V, the encryption algorithm is Las Vegas and may require arbitrarily many random bits (although with extremely small probability). It is possible to modify this scheme for limiting the number of requested random bits, but this modification is cumbersome. The principle is as follows. As the distribution of the required number of random bits has a very small tail (similar to a Gaussian distribution), one could fix an *a priori* bound to the number of requested bits and abort the algorithm and return an arbitrary value if that bound is reached. Since the output distributions of the original and modified encryption algorithms are statistically very close, properties that hold for the original scheme still hold for the modified scheme.

B. Strong Secrecy and Semantic Security

In what follows we will consider both continuous and discrete random variables as well as *mixed pairs* of discrete and continuous random variables. Let X, Y be continuous random variables taking values in $\mathbb{R}^n$ with densities p_{X} and p_{Y} respectively, and $\mathsf{M}, \bar{\mathsf{M}}$ discrete random variables taking values in a finite set $\mathcal{M}_n$, with probability mass functions $p_{\mathsf{M}}, p_{\bar{\mathsf{M}}}$. Let $p_{\mathsf{XM}}(\mathbf{x}, m)$ be the *joint hybrid density* of the mixed pair (X, M): that is, $\forall m \in \mathcal{M}_n$, $p_{\mathsf{XM}}(\cdot, m)$ is the density corresponding to the probability measure $\mu_m(A) = \mathbb{P}\{\mathsf{M} = m, \mathsf{X} \in A\}$ for all measurable sets $A \subseteq \mathbb{R}^n$.

Definition 2 (Kullback-Leibler divergence and mutual information)**.** *The Kullback-Leibler divergence of the continuous distributions p_{X} and p_{Y} is defined as $\mathbb{D}(p_{\mathsf{X}} \| p_{\mathsf{Y}}) = \int_{\mathbb{R}^n} p_{\mathsf{X}}(\mathbf{x}) \log \frac{p_{\mathsf{X}}(\mathbf{x})}{p_{\mathsf{Y}}(\mathbf{x})} d\mathbf{x}$. Similarly, for discrete distributions p_{M} and $p_{\bar{\mathsf{M}}}$ we define $\mathbb{D}(p_{\mathsf{M}} \| p_{\bar{\mathsf{M}}}) = \sum_{m \in \mathcal{M}_n} p_{\mathsf{M}}(m) \log \frac{p_{\mathsf{M}}(m)}{p_{\bar{\mathsf{M}}}(m)}$. The mutual information between X, Y is defined by*

$$\mathbb{I}(\mathsf{X}; \mathsf{Y}) = \mathbb{D}(p_{\mathsf{XY}} \| p_{\mathsf{X}} p_{\mathsf{Y}}).$$

We now recall the notion of *variational distance* or *statistical distance* between two distributions. Note that in the literature the variational distance is sometimes scaled by a factor $\frac{1}{2}$. We choose normalization factor 1 so that it matches with the L^1 distance between distributions.

Definition 3 (Variational distance)**.** *Let p and q be two discrete distributions on a finite set $\mathcal{M}$. Then the variational distance between p and q is*

$$\mathbb{V}(p, q) \triangleq \sum_{x \in \mathcal{X}} |p(x) - q(x)| .$$

Similarly, for continuous distributions with densities p and q, the variational distance is defined by

$$\mathbb{V}(p, q) \triangleq \int |p(x) - q(x)| \, dx.$$

With the definitions given above, we are ready to introduce strong secrecy and semantic security.

Definition 4 (Achievable strong secrecy rate)**.** *The message rate R is an* achievable strong secrecy rate *if there exists a sequence of wiretap codes $\{\mathcal{C}_n\}$ of rate R such that*

$$\mathbb{P}\{\hat{\mathsf{M}} \neq \mathsf{M}\} \to 0, \qquad \text{(reliability)}$$
$$\mathbb{I}(\mathsf{M}; \mathsf{Z}^n) \to 0 \qquad \text{(strong secrecy)}$$

when $n \to \infty$.

In the definition of strong secrecy for communications, no special attention is paid to the issue of message distribution. In fact, a uniform distribution is often assumed in the coding literature. But this is insufficient from a cryptographic viewpoint, as it does not ensure security for a particular message. To address this issue of the wiretap code, we need to ensure the mutual information vanishes for all message distributions:

$$\mathrm{Adv}^{\mathrm{mis}}(\mathsf{Z}^n) \triangleq \max_{p_{\mathsf{M}}} \mathbb{I}(\mathsf{M}; \mathsf{Z}^n) \to 0 \tag{3}$$

when $n \to \infty$. The *adversarial advantage* $\mathrm{Adv}^{\mathrm{mis}}$ tending to zero was termed *mutual information security* in [4]. In this paper, the terms mutual information security and strong secrecy for all message distributions are used interchangeably. Note that one may further impose constraints on the rate of convergence towards 0; in practice an exponential rate of convergence is desired.

Let the min-entropy of a discrete random variable M be

$$\mathbb{H}_\infty(\mathsf{M}) = -\log\left(\max_m \mathbb{P}\{\mathsf{M} = m\}\right),$$

and the conditional min-entropy of M given U be

$$\mathbb{H}_\infty(\mathsf{M}|\mathsf{U}) = \sum_u \mathbb{P}\{\mathsf{U} = u\}\mathbb{H}_\infty(\mathsf{M}|\mathsf{U} = u).$$

Definition 5 (Semantic security)**.** *A sequence of wiretap codes $\{\mathcal{C}_n\}$ achieves semantic security if*

$$\mathrm{Adv}^{\mathrm{ss}}(\mathsf{Z}^n) \triangleq \sup_{f,p_{\mathsf{M}}} \left(e^{-\mathbb{H}_\infty(f(\mathsf{M})|\mathsf{Z}^n)} - e^{-\mathbb{H}_\infty(f(\mathsf{M}))} \right) \to 0$$

when $n \to \infty$. The supremum is taken over all message distributions p_{M} and all functions f of M taking values in the set $\{0,1\}^*$ of finite binary words.

Semantic security means that, asymptotically, it is impossible to estimate any function of the message better than to guess it without considering Z^n at all. We also define distinguishing security, which means that, asymptotically, the channel outputs are indistinguishable for different input messages.

Definition 6 (Distinguishing security)**.** *A sequence of wiretap codes $\{\mathcal{C}_n\}$ achieves distinguishing security if*

$$\mathrm{Adv}^{\mathrm{ds}}(\mathsf{Z}^n) \triangleq \max_{m,m'} \mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, p_{\mathsf{Z}^n|\mathsf{M}=m'}) \to 0, \tag{4}$$

when $n \to \infty$. The maximum in the previous equation is taken over all messages $m, m' \in \mathcal{M}_n$.

As for the discrete wiretap channel setup considered in [4], the classical proof of equivalence between semantic security and distinguishing security [3] can be readily adapted and it can be shown that[1]

$$2\mathrm{Adv}^{\mathrm{ss}}(\mathsf{Z}^n) \le \mathrm{Adv}^{\mathrm{ds}}(\mathsf{Z}^n) \le 4\mathrm{Adv}^{\mathrm{ss}}(\mathsf{Z}^n). \tag{5}$$

Even though the two definitions are equivalent, distinguishing security often turns out to be technically easier to manipulate.

C. Equivalence

We will show that semantic security and strong secrecy for all message distributions are equivalent for continuous channels. This is an extension of the results from Section 3 of [4].

We first need the following continuous channel adaptation of Csiszár's in [2, Lemma 1]. The lower bound is a consequence of Pinsker's inequality (see [22, pp.58-59]). The proof of the upper bound is similar to the discrete case and is given in Appendix I.

Lemma 1. *Let Z^n be a random variable defined on $\mathbb{R}^n$ and M be a random variable over a finite domain $\mathcal{M}_n$ such that $|\mathcal{M}_n| \ge 4$. Then*

$$\frac{1}{2}d_{\mathrm{av}}^2 \le \mathbb{I}(\mathsf{M};\mathsf{Z}^n) \le d_{\mathrm{av}} \log \frac{|\mathcal{M}_n|}{d_{\mathrm{av}}},$$

where

$$d_{\mathrm{av}} = \sum_{m\in\mathcal{M}_n} p_{\mathsf{M}}(m)\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, p_{\mathsf{Z}^n})$$

is the average variational distance of the conditional output distributions from the global output distribution.

We now prove the equivalence between semantic security and strong secrecy for all message distributions via distinguishing security.

Theorem 1. *a) A sequence of wiretap codes $\{\mathcal{C}_n\}$ of rate R which achieves semantic security with advantage $\mathrm{Adv}^{\mathrm{ds}}(\mathsf{Z}^n) = o\left(\frac{1}{n}\right)$ also achieves strong secrecy for all message distributions: $\forall p_{\mathsf{M}}$,*

$$\mathbb{I}(\mathsf{M};\mathsf{Z}^n) \le \mathrm{Adv}^{\mathrm{mis}}(\mathsf{Z}^n) \le \varepsilon_n \left(nR - \log \varepsilon_n\right),$$

where $\varepsilon_n \triangleq \mathrm{Adv}^{\mathrm{ds}}(\mathsf{Z}^n)$. b) A sequence of wiretap codes $\{\mathcal{C}_n\}$ which achieves strong secrecy for all message distributions also achieves semantic security:

$$\mathrm{Adv}^{\mathrm{ds}}(\mathsf{Z}^n) \le 2\sqrt{2\mathrm{Adv}^{\mathrm{mis}}(\mathsf{Z}^n)}.$$

Proof:

(a) Distinguishing security ⇒ strong secrecy for all message distributions: For any $m \in \mathcal{M}_n$, we have

$$\begin{aligned}
&\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, p_{\mathsf{Z}^n}) \\
&= \int_{\mathbb{R}^n} \Big| p_{\mathsf{Z}^n|\mathsf{M}}(\mathbf{z}|m) - \sum_{m'\in\mathcal{M}_n} p_{\mathsf{M}}(m')p_{\mathsf{Z}^n|\mathsf{M}}(\mathbf{z}|m') \Big| d\mathbf{z} \\
&= \int_{\mathbb{R}^n} \Big| \sum_{m'\in\mathcal{M}_n} p_{\mathsf{M}}(m') \left(p_{\mathsf{Z}^n|\mathsf{M}}(\mathbf{z}|m) - p_{\mathsf{Z}^n|\mathsf{M}}(\mathbf{z}|m') \right) \Big| d\mathbf{z} \\
&\le \max_{m'\in\mathcal{M}_n} \mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, p_{\mathsf{Z}^n|\mathsf{M}=m'}) \\
&\le \max_{m',m''\in\mathcal{M}_n} \mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m'}, p_{\mathsf{Z}^n|\mathsf{M}=m''}) = \varepsilon_n.
\end{aligned}$$

Therefore $d_{\mathrm{av}} \le \varepsilon_n$. By Lemma 1, we obtain

$$\mathbb{I}(\mathsf{M};\mathsf{Z}^n) \le \varepsilon_n \log \frac{|\mathcal{M}_n|}{\varepsilon_n} = \varepsilon_n nR - \varepsilon_n \log \varepsilon_n.$$

If $\mathrm{Adv}^{\mathrm{ds}}(\mathsf{Z}^n) = o(\frac{1}{n})$, then $\mathbb{I}(\mathsf{M};\mathsf{Z}^n) \to 0$.

(b) Strong secrecy for all message distributions ⇒ distinguishing security: Let $m \in \mathcal{M}_n$ be arbitrary. If strong secrecy holds for all distributions, then in particular it holds for the distribution p_m defined by $p_m(m') = 1$ if $m = m'$ and 0 otherwise. Now, Pinsker's inequality (see [22, pp.58-59]) asserts that $\mathbb{V}(p,q) \le \sqrt{2\mathbb{D}(p\|q)}$ for any distributions p and q. We thus have:

$$\begin{aligned}
&\mathbb{V}(p_{(\mathsf{Z}^n,m)}, p_{\mathsf{Z}^n}p_m) \\
&\qquad = \sum_{m'} \int_{\mathbb{R}^n} \left| p_{(\mathsf{Z}^n,m)}(\mathbf{z},m') - p_{\mathsf{Z}^n}(\mathbf{z})p_m(m') \right| d\mathbf{z} \\
&\qquad = \int_{\mathbb{R}^n} \left| p_{\mathsf{Z}^n|\mathsf{M}=m}(\mathbf{z}) - p_{\mathsf{Z}^n}(\mathbf{z}) \right| d\mathbf{z} \\
&\qquad \le \sqrt{2\mathbb{I}(m;\mathsf{Z}^n)}.
\end{aligned}$$

The strong secrecy assumption implies that:

$$\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, p_{\mathsf{Z}^n}) = \int_{\mathbb{R}^n} \left| p_{\mathsf{Z}^n|\mathsf{M}=m}(\mathbf{z}) - p_{\mathsf{Z}^n}(\mathbf{z}) \right| d\mathbf{z} \to 0.$$

Using the triangular inequality

$$\begin{aligned}
&\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, p_{\mathsf{Z}^n|\mathsf{M}=m'}) \\
&\qquad \le \mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, p_{\mathsf{Z}^n}) + \mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m'}, p_{\mathsf{Z}^n}),
\end{aligned}$$

we obtain distinguishing security. □

Note that Lemma 2 in [2] also holds: For any distribution q_{Z^n} on $\mathbb{R}^n$, we have

$$d_{\mathrm{av}} \le 2\sum_{m\in\mathcal{M}_n} p_{\mathsf{M}}(m)\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, q_{\mathsf{Z}^n}). \tag{6}$$

[1] Note that the factors in [4] are 1 on the left and 2 on the right, respectively, due to the factor $\frac{1}{2}$ used in the definition of the variational distance in [4].

Together with Lemma 1, this leads to an upper bound on the mutual information, in case we can approximate $p_{\mathsf{Z}^n|\mathsf{M}=m}$ by a density that is independent of m.

Lemma 2. *Suppose that for all n there exists some density q_{Z^n} in $\mathbb{R}^n$ such that $\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, q_{\mathsf{Z}^n}) \leq \varepsilon_n$, for all $m \in \mathcal{M}_n$. Then we have $d_{\text{av}} \leq 2\varepsilon_n$ and so*

$$\mathbb{I}(\mathsf{M}; \mathsf{Z}^n) \leq 2\varepsilon_n nR - 2\varepsilon_n \log(2\varepsilon_n). \tag{7}$$

In the rest of this paper, we will use lattice codes to achieve semantic security.

III. Lattice Gaussian distribution and flatness factor

In this section, we introduce the mathematical tools we will need to describe and analyze our wiretap codes.

A. Preliminaries on Lattices

An n-dimensional lattice Λ in the Euclidean space $\mathbb{R}^n$ is a set defined by

$$\Lambda = \mathcal{L}(\mathbf{B}) = \{\mathbf{B}\mathbf{x} : \mathbf{x} \in \mathbb{Z}^n\}$$

where the columns of the basis matrix $\mathbf{B} = [\mathbf{b}_1 \cdots \mathbf{b}_n]$ are linearly independent. (In this work, we will restrict ourselves to full-rank lattices.) The dual lattice Λ^* of a lattice Λ is defined as the set of vectors $\mathbf{v} \in \mathbb{R}^n$ such that $\langle \mathbf{v}, \lambda \rangle \in \mathbb{Z}$, for all $\lambda \in \Lambda$ (see, e.g., [23]).

For a vector $\mathbf{x}$, the nearest-neighbor quantizer associated with Λ is $Q_\Lambda(\mathbf{x}) = \arg\min_{\lambda \in \Lambda} \|\lambda - \mathbf{x}\|$. We define the modulo lattice operation by $\mathbf{x} \mod \Lambda \triangleq \mathbf{x} - Q_\Lambda(\mathbf{x})$. The Voronoi cell of Λ, defined by $\mathcal{V}(\Lambda) = \{\mathbf{x} : Q_\Lambda(\mathbf{x}) = \mathbf{0}\}$, specifies the nearest-neighbor decoding region. Important quantities for $\mathcal{V}(\Lambda)$ include the cell volume $V(\Lambda) = \int_{\mathcal{V}(\Lambda)} d\mathbf{x}$, the second moment per dimension $\sigma^2(\Lambda) = \frac{1}{nV(\Lambda)} \int_{\mathcal{V}(\Lambda)} \|\mathbf{x}\|^2 d\mathbf{x}$ and the normalized second moment $G(\Lambda) = \sigma^2(\Lambda)/V(\Lambda)^{\frac{2}{n}}$. The minimum of $G(\Lambda)$ of all the n-dimensional lattices is denoted as G_n. From [24, p.58], we have $G_n \geq 1/(2\pi e)$ for all n, and $\lim_{n\to\infty} G_n = 1/(2\pi e)$. The Voronoi cell is one example of fundamental region of the lattice. A measurable set $\mathcal{R} \subset \mathbb{R}^n$ is a fundamental region of the lattice Λ if $\cup_{\lambda \in \Lambda}(\mathcal{R} + \lambda) = \mathbb{R}^n$ and if $(\mathcal{R} + \lambda) \cap (\mathcal{R} + \lambda')$ has measure 0 for any $\lambda \neq \lambda'$ in Λ. The volume of a fundamental region is equal to that of the Voronoi cell $V(\Lambda)$.

For a (full-rank) sublattice $\Lambda' \subset \Lambda$, the finite group Λ/Λ' is defined as the group of distinct cosets $\lambda + \Lambda'$ for $\lambda \in \Lambda$. The lattices Λ' and Λ are often said to form a pair of nested lattices, in which Λ is referred to as the fine lattice while Λ' the coarse lattice. The nesting ratio is equal to $V(\Lambda')/V(\Lambda)$.

Some background on lattices that are good for channel coding, and have been shown to approach the capacity of the Gaussian channel, is provided in Appendix II. This includes definitions of Rogers, quantization and AWGN-goodness.

B. Lattice Theta Series

The theta series of Λ (see, e.g., [24]) is defined as

$$\Theta_\Lambda(q) = \sum_{\lambda \in \Lambda} q^{\|\lambda\|^2} \tag{8}$$

where $q = e^{j\pi z}$ ($\Im(z) > 0$). Letting z be purely imaginary, and assuming $\tau = \Im(z) > 0$, we can alternatively express the theta series as

$$\Theta_\Lambda(\tau) = \sum_{\lambda \in \Lambda} e^{-\pi\tau\|\lambda\|^2}. \tag{9}$$

In [13], Loeliger derived a version of the Minkowski-Hlawka theorem based on the averaging over Construction-A lattices. We adapt his method to derive the average behavior of the theta series for Construction A. Loeliger's derivation has a restriction in that it requires a function of bounded support, which is not the case for the Gaussian function associated with the theta series. This restriction is circumvented here.

For integer $p > 0$, let $\mathbb{Z}^n \to \mathbb{Z}_p^n : \mathbf{v} \mapsto \overline{\mathbf{v}}$ be the element-wise reduction modulo-p. Following [13], consider mod-p lattices (Construction A) of the form $\Lambda_C \triangleq \{\mathbf{v} \in \mathbb{Z}^n : \overline{\mathbf{v}} \in C\}$, where p is a prime and C is a linear code over $\mathbb{Z}_p$. In the proof, scaled mod-p lattices $a\Lambda_C \triangleq \{a\mathbf{v} : \mathbf{v} \in \Lambda_C\}$ for some $a \in \mathbb{R}^+$ are used. The fundamental volume of such a lattice is $V(a\Lambda_C) = a^n p^{n-k}$, where n and k are the block length and dimension of the code C, respectively. A set $\mathcal{C}$ of linear codes over $\mathbb{Z}_p$ is said to be balanced if every nonzero element of $\mathbb{Z}_p^n$ is contained in the same number of codes from $\mathcal{C}$. In particular, the set of all linear (n, k) codes over $\mathbb{Z}_p$ is balanced.

Lemma 3 (Average behavior of theta series)**.** *Let $\mathcal{C}$ be any balanced set of linear (n, k) codes over $\mathbb{Z}_p$. Then, for $0 < k < n$, for $a^n p^{n-k} = V$ and τ fixed, we have:*

$$\lim_{a\to 0, p\to\infty} \frac{1}{|\mathcal{C}|} \sum_{C \in \mathcal{C}} \Theta_{a\Lambda_C}(\tau) = 1 + \frac{1}{V\tau^{n/2}}. \tag{10}$$

The proof of Lemma 3 is provided in Appendix IV-A.

C. Lattice Gaussian Distribution

Lattice Gaussian distributions arise from various problems in mathematics [25], coding [14] and cryptography [26]. For $\sigma > 0$ and $\mathbf{c} \in \mathbb{R}^n$, we define the Gaussian distribution of variance σ centered at $\mathbf{c} \in \mathbb{R}^n$ as

$$f_{\sigma,\mathbf{c}}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\|\mathbf{x}-\mathbf{c}\|^2}{2\sigma^2}},$$

for all $\mathbf{x} \in \mathbb{R}^n$. For convenience, we write $f_\sigma(\mathbf{x}) = f_{\sigma,\mathbf{0}}(\mathbf{x})$.

We also consider the Λ-periodic function

$$f_{\sigma,\Lambda}(\mathbf{x}) = \sum_{\lambda \in \Lambda} f_{\sigma,\lambda}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\lambda \in \Lambda} e^{-\frac{\|\mathbf{x}-\lambda\|^2}{2\sigma^2}}, \tag{11}$$

for all $\mathbf{x} \in \mathbb{R}^n$. Observe that $f_{\sigma,\Lambda}$ restricted to the quotient $\mathbb{R}^n/\Lambda$ is a probability density.

We define the *discrete Gaussian distribution* over Λ centered at $\mathbf{c} \in \mathbb{R}^n$ as the following discrete distribution taking values in $\lambda \in \Lambda$:

$$D_{\Lambda,\sigma,\mathbf{c}}(\lambda) = \frac{f_{\sigma,\mathbf{c}}(\lambda)}{f_{\sigma,\mathbf{c}}(\Lambda)} \quad \forall \lambda \in \Lambda,$$

where $f_{\sigma,\mathbf{c}}(\Lambda) \triangleq \sum_{\lambda\in\Lambda} f_{\sigma,\mathbf{c}}(\lambda)$. Again for convenience, we write $D_{\Lambda,\sigma} = D_{\Lambda,\sigma,\mathbf{0}}$. We remark that this definition differs slightly from the one in [26], where σ is scaled by a constant factor $\sqrt{2\pi}$ (i.e., $s = \sqrt{2\pi}\sigma$).

It will be useful to define the discrete Gaussian distribution over a coset of Λ, i.e., the shifted lattice $\Lambda - \mathbf{c}$:

$$D_{\Lambda-\mathbf{c},\sigma}(\lambda - \mathbf{c}) = \frac{f_\sigma(\lambda - \mathbf{c})}{f_{\sigma,\mathbf{c}}(\Lambda)} \quad \forall \lambda \in \Lambda.$$

Note the relation $D_{\Lambda-\mathbf{c},\sigma}(\lambda - \mathbf{c}) = D_{\Lambda,\sigma,\mathbf{c}}(\lambda)$, namely, they are a shifted version of each other.

D. Flatness Factor

The flatness factor of a lattice Λ quantifies the maximum variation of $f_{\sigma,\Lambda}(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^n$.

Definition 7 (Flatness factor). *For a lattice Λ and for a parameter σ, the flatness factor is defined by:*

$$\epsilon_\Lambda(\sigma) \triangleq \frac{\max_{\mathbf{x}\in\mathcal{R}(\Lambda)} |f_{\sigma,\Lambda}(\mathbf{x}) - 1/V(\Lambda)|}{1/V(\Lambda)}.$$

In other words, $f_{\sigma,\Lambda}(\mathbf{x})$ is within $1 \pm \epsilon_\Lambda(\sigma)$ from the uniform distribution over $\mathcal{R}(\Lambda)$. The flatness factor may also be interpreted as a scaled maximum variation of $f_{\sigma,\Lambda}(\mathbf{x})$, as $\mathbb{E}_\mathbf{x}[f_{\sigma,\Lambda}(\mathbf{x})] = 1/V(\Lambda)$ when $\mathbf{x}$ is sampled uniformly in $\mathcal{R}(\Lambda)$. Note that this definition slightly differs from that in [27]: The present definition also takes into account the minimum of $f_{\sigma,\Lambda}(\mathbf{x})$.

Proposition 1 (Expression of $\epsilon_\Lambda(\sigma)$). *We have:*

$$\epsilon_\Lambda(\sigma) = \gamma_\Lambda(\sigma)^{\frac{n}{2}} \Theta_\Lambda\left(\frac{1}{2\pi\sigma^2}\right) - 1$$

where $\gamma_\Lambda(\sigma) = \frac{V(\Lambda)^{\frac{2}{n}}}{2\pi\sigma^2}$ is the generalized signal-to-noise ratio (GSNR)[2].

Proof: Using the Fourier expansion of $f_{\sigma,\Lambda}(\mathbf{x})$ (see, e.g., [14, 26]), we obtain, for all $\mathbf{x} \in \mathcal{R}(\Lambda)$:

$$\begin{aligned}
V(\Lambda)&\left|f_{\sigma,\Lambda}(\mathbf{x}) - \frac{1}{V(\Lambda)}\right| \\
&= \left|\sum_{\lambda^*\in\Lambda^*} e^{-2\pi^2\sigma^2\|\lambda^*\|^2} \cos(2\pi\langle\lambda^*, \mathbf{x}\rangle) - 1\right| \\
&\overset{(a)}{\leq} \sum_{\lambda^*\in\Lambda^*} e^{-2\pi^2\sigma^2\|\lambda^*\|^2} - 1 \\
&\overset{(b)}{=} V(\Lambda) f_{\sigma,\Lambda}(\mathbf{0}) - 1 \\
&= \frac{V(\Lambda)}{(\sqrt{2\pi}\sigma)^n} \sum_{\lambda\in\Lambda} e^{-\frac{\|\lambda\|^2}{2\sigma^2}} - 1 \\
&\overset{(c)}{=} \frac{V(\Lambda)}{(\sqrt{2\pi}\sigma)^n} \Theta_\Lambda\left(\frac{1}{2\pi\sigma^2}\right) - 1,
\end{aligned}$$

where the equality in (a) holds if $\mathbf{x} \in \Lambda$ so that $\langle\lambda^*, \mathbf{x}\rangle$ is an integer for all $\lambda^* \in \Lambda^*$, (b) is due to the Poisson sum formula, and (c) follows from the definition of the theta series. The result follows. □

Remark 2. The equality in (a) implies that the maxima of both $f_{\sigma,\Lambda}(\mathbf{x})$ and $|f_{\sigma,\Lambda}(\mathbf{x}) - 1/V(\Lambda)|$ are reached when $\mathbf{x} \in \Lambda$.

Remark 3. From the expression

$$\epsilon_\Lambda(\sigma) = \sum_{\lambda^*\in\Lambda^*} e^{-2\pi^2\sigma^2\|\lambda^*\|^2} - 1,$$

it is easy to see that ϵ_Λ is a monotonically decreasing function, i.e., for $\sigma_1 < \sigma_2$, we have $\epsilon_\Lambda(\sigma_2) \leq \epsilon_\Lambda(\sigma_1)$.

Remark 4. If Λ_2 is a sublattice of Λ_1, then $\epsilon_{\Lambda_1}(\sigma) \leq \epsilon_{\Lambda_2}(\sigma)$.

Remark 5. The flatness factor is scaling invariant, i.e., $\epsilon_\Lambda(\sigma) = \epsilon_{a\Lambda}(a\sigma)$.

In the following, we show that the flatness factor is equivalent to the notion of smoothing parameter that is commonly used in lattice-based cryptography.

Definition 8 (Smoothing parameter [26]). *For a lattice Λ and for $\varepsilon > 0$, the smoothing parameter $\eta_\varepsilon(\Lambda)$ is the smallest $s > 0$ such that $\sum_{\lambda^*\in\Lambda^*\setminus\{\mathbf{0}\}} e^{-\pi s^2\|\lambda^*\|^2} \leq \varepsilon$.*

Proposition 2. *If $\eta_\varepsilon(\Lambda) = \sqrt{2\pi}\sigma$, then $\epsilon_\Lambda(\sigma) = \varepsilon$.*

Proof: From the proof of Proposition 1, we can see that

$$\epsilon_\Lambda(\sigma) = \sum_{\lambda^*\in\Lambda^*} e^{-2\pi^2\sigma^2\|\lambda^*\|^2} - 1 = \sum_{\lambda^*\in\Lambda^*\setminus\{\mathbf{0}\}} e^{-\pi s^2\|\lambda^*\|^2} = \varepsilon$$

for $s = \sqrt{2\pi}\sigma$. □

Despite the equivalence, the flatness factor has two main technical advantages:

- It allows for a direct characterization by the theta series. Note that it is ε, not the smoothing parameter, that is of more interest to communications.
- The studies of the smoothing parameter are mostly concerned with small values of ε, while the flatness factor can handle both large and small values of ε. This is of interest in communication applications [27].

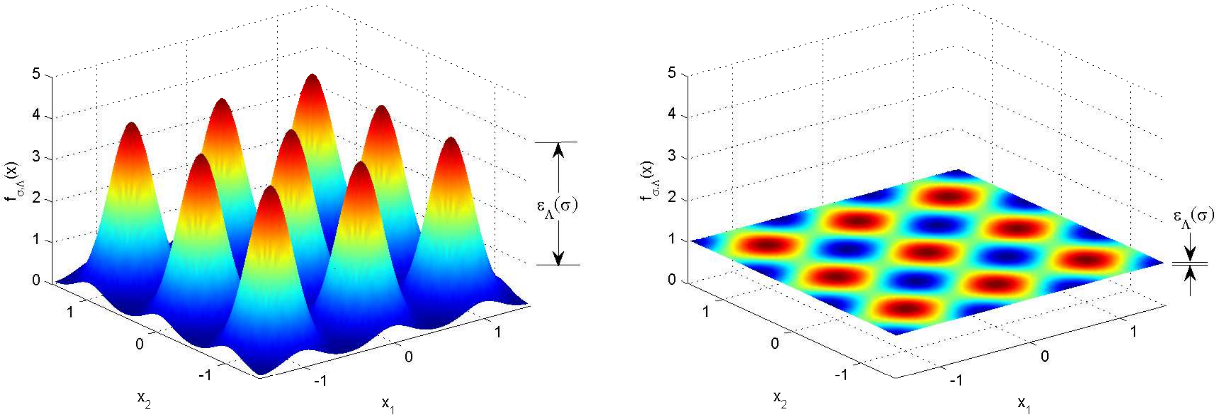
Figure 1 illustrates the flatness factor and lattice Gaussian distribution at different GSNRs for lattice $\mathbb{Z}^2$. When the GSNR is high (Fig. 1(a)), $\epsilon_\Lambda(\sigma)$ is large and the Gaussians are well separated, implying reliable decoding is possible; this scenario is desired in communications. When the GSNR is low (Fig. 1(b)), $\epsilon_\Lambda(\sigma)$ is small and the distribution is nearly uniform, implying reliable decoding is impossible; this scenario is desired in security and will be pursued in following sections.

The flatness factor also gives a bound on the variational distance between the Gaussian distribution reduced mod $\mathcal{R}(\Lambda)$ and the uniform distribution $U_{\mathcal{R}(\Lambda)}$ on $\mathcal{R}(\Lambda)$. This result was proven in [26] using the smoothing parameter when $\mathcal{R}(\Lambda)$ is the fundamental parallelotope. We give a proof for any $\mathcal{R}(\Lambda)$, for the sake of completeness.

Proposition 3. *For $\mathbf{c} \in \mathbb{R}^n$, let $\bar{f}(\mathbf{x})$ be the density function of the distribution over $\mathcal{R}(\Lambda)$ defined by $f_{\sigma,\mathbf{c}} \bmod \mathcal{R}(\Lambda)$. Then*

$$\mathbb{V}(\bar{f}, U_{\mathcal{R}(\Lambda)}) \leq \epsilon_\Lambda(\sigma).$$

[2] Note that this definition of GSNR is slightly different from similar definitions in literature, by a factor 2π or e. In particular, Poltyrev defined the GSNR as $V(\Lambda)^{\frac{2}{n}}/\sigma^2$ [28], while the volume-to-noise ratio (VNR) is defined as $V(\Lambda)^{\frac{2}{n}}/(2\pi e\sigma^2)$ in [14, 15].



(a) $\gamma_\Lambda(\sigma) = 4, \epsilon_\Lambda(\sigma) = 3$.

(b) $\gamma_\Lambda(\sigma) = 0.5, \epsilon_\Lambda(\sigma) = 0.0075$.

Fig. 1. Lattice Gaussian distribution and flatness factor for $\mathbb{Z}^2$ (a) at high GSNR where $\epsilon_\Lambda(\sigma)$ is large and the Gaussians are well separated, and (b) at low GSNR where $\epsilon_\Lambda(\sigma)$ is small and the distribution is nearly uniform.

Proof: Observe that restricting $f_{\sigma,\Lambda}$ to any fundamental region $\mathcal{R}(\Lambda)$ is equivalent to considering the Gaussian distribution modulo $\mathcal{R}(\Lambda)$:

$$\begin{aligned}\bar{f}(\mathbf{x}) &= \sum_{\lambda\in\Lambda} f_{\sigma,\mathbf{c}}(\mathbf{x}-\lambda)\mathbb{1}_{\mathcal{R}(\Lambda)}(\mathbf{x}) \\ &= \sum_{\lambda\in\Lambda} f_{\sigma,\lambda}(\mathbf{x}-\mathbf{c})\mathbb{1}_{\mathcal{R}(\Lambda)}(\mathbf{x}) \\ &= f_{\sigma,\Lambda}(\mathbf{x}-\mathbf{c})\mathbb{1}_{\mathcal{R}(\Lambda)}(\mathbf{x}).\end{aligned}$$

Then by definition of $\epsilon_\Lambda(\sigma)$, we find

$$\begin{aligned}&\int_{\mathcal{R}(\Lambda)} \left|\bar{f}(\mathbf{t}) - U_{\mathcal{R}(\Lambda)}(\mathbf{t})\right| d\mathbf{t} \\ &\quad\le V(\Lambda) \max_{\mathbf{x}\in\mathcal{R}(\Lambda)} \left| f_{\sigma,\Lambda}(\mathbf{x}-\mathbf{c}) - \frac{1}{V(\Lambda)}\right| \\ &\quad= V(\Lambda) \max_{\mathbf{x}\in\mathcal{R}(\Lambda)-\mathbf{c}} \left| f_{\sigma,\Lambda}(\mathbf{x}) - \frac{1}{V(\Lambda)}\right| \le \epsilon_\Lambda(\sigma),\end{aligned}$$

because $\mathcal{R}(\Lambda) - \mathbf{c}$ is a fundamental region of Λ. □

By definition, the flatness factor in fact guarantees a stronger property: if $\epsilon_\Lambda(\sigma) \to 0$, then $f_{\sigma,\Lambda}(\mathbf{x})$ converges uniformly to the uniform distribution on the fundamental region.

The following result guarantees the existence of sequences of lattices whose flatness factors can respectively vanish or explode as $n \to \infty$.

Theorem 2. *$\forall\sigma > 0$ and $\forall\delta > 0$, there exists a sequence of mod-p lattices $\Lambda^{(n)}$ such that*

$$\epsilon_{\Lambda^{(n)}}(\sigma) \le (1+\delta)\cdot\gamma_{\Lambda^{(n)}}(\sigma)^{\frac{n}{2}}, \tag{12}$$

i.e., the flatness factor goes to zero exponentially as long as the GSNR satisfies $\gamma_{\Lambda^{(n)}}(\sigma) < 1$; oppositely, there also exists a sequence of mod-p lattices $\Lambda'^{(n)}$ such that

$$\epsilon_{\Lambda'^{(n)}}(\sigma) \ge (1-\delta)\cdot\gamma_{\Lambda'^{(n)}}(\sigma)^{\frac{n}{2}}, \tag{13}$$

i.e., its flatness factor goes to infinity exponentially as long as $\gamma_{\Lambda'^{(n)}}(\sigma) > 1$.

Proof: Lemma 3 guarantees that for all n, δ and τ there exists $a(n,\delta,\tau)$ (and the corresponding p such that $a^n p^{n-k} = V(\Lambda)$) such that $\mathbb{E}_C\left[\Theta_{a\Lambda_C}(\tau)\right] \le 1+\delta+\frac{1}{V(\Lambda)\tau^{\frac{n}{2}}}$. Here C is sampled uniformly among all linear (n,k) codes over $\mathbb{Z}_p$ and $a\Lambda_C = \{a\mathbf{v} : \mathbf{v}\in\Lambda_C\}$. Therefore there exists a sequence of lattices $\Lambda^{(n)}$ such that $\Theta_{\Lambda^{(n)}}(\tau) \le 1+\delta+\frac{1}{V(\Lambda^{(n)})\tau^{\frac{n}{2}}}$. For this sequence, Proposition 1 gives $\epsilon_\Lambda(\sigma) \le (1+\delta)\gamma_\Lambda(\sigma)^{\frac{n}{2}}$ when we let $\tau = \frac{1}{2\pi\sigma^2}$. The second half of the theorem can be proved in a similar fashion. □

Theorem 2 shows a phenomenon of "phase transition" for the flatness factor, where the boundary is $\gamma_\Lambda(\sigma) = 1$.

Remark 6. In fact, we can show a concentration result on the flatness factor of the ensemble of mod-p lattices, that is, most mod-p lattices have a flatness factor concentrating around $\gamma_{\Lambda^{(n)}}(\sigma)^{\frac{n}{2}}$. In particular, using the Markov inequality, we see that with probability higher than $1-2^{-n}$ over the choice of $\Lambda^{(n)}$,

$$\epsilon_{\Lambda^{(n)}}(\sigma) \le (1+\delta)\cdot\left[4\gamma_{\Lambda^{(n)}}(\sigma)\right]^{\frac{n}{2}}, \tag{14}$$

Thus, for $\gamma_{\Lambda^{(n)}}(\sigma) < 1/4$, we could have $\epsilon_\Lambda(\sigma) \to 0$ exponentially. This is slightly worse than what we have in (12), but it holds with very high probability, making the construction of the scheme potentially more practical.

E. Properties of the Flatness Factor

In this section we collect known properties and further derive new properties of lattice Gaussian distributions that will be useful in the paper.

From the definition of the flatness factor and Remark 2, one can derive the following result (see also [26, Lemma 4.4]):

Lemma 4. *For all $\mathbf{c}\in\mathbb{R}^n$ and $\sigma > 0$, we have:*

$$\frac{f_{\sigma,\mathbf{c}}(\Lambda)}{f_\sigma(\Lambda)} \in \left[\frac{1-\epsilon_\Lambda(\sigma)}{1+\epsilon_\Lambda(\sigma)}, 1\right].$$

The following lemma shows that, when the flatness factor of the coarse lattice is small, a discrete Gaussian distribution over the fine lattice results in almost uniformly distributed cosets, and vice versa. The first half of the lemma is a corollary of Lemma 4 (see [29, Corollary 2.7]), while the second half is proven in Appendix IV-B.

Lemma 5. *Let $\Lambda' \subset \Lambda$ be a pair of nested lattices such that $\epsilon_{\Lambda'}(\sigma) < \frac{1}{2}$. Then*

$$\mathbb{V}(D_{\Lambda,\sigma,\mathbf{c}} \bmod \Lambda', U(\Lambda/\Lambda')) \le 4\epsilon_{\Lambda'}(\sigma),$$

where $U(\Lambda/\Lambda')$ denotes the uniform distribution over the finite set Λ/Λ'. Conversely, if L is uniformly distributed in Λ/Λ' and L' is sampled from $D_{\Lambda',\sigma,\mathsf{L}}$, then

$$\mathbb{V}(\mathsf{L}+\mathsf{L}', D_{\Lambda,\sigma}) \le \frac{2\epsilon_{\Lambda'}(\sigma)}{1-\epsilon_{\Lambda'}(\sigma)}.$$

The following result shows that the variance per dimension of the discrete Gaussian $D_{\Lambda,\sigma,\mathbf{c}}$ is not too far from σ^2 when the flatness factor is small. The result follows easily by combining Lemma 4.2 and the proof of Lemma 4.3 in [26].

Lemma 6. *Let L be sampled from the Gaussian distribution $D_{\Lambda,\sigma,\mathbf{c}}$. If $\varepsilon \triangleq \epsilon_\Lambda(\sigma/2) < 1$, then*

$$\left|\mathbb{E}\left[\|\mathsf{L}-\mathbf{c}\|^2\right] - n\sigma^2\right| \le \frac{2\pi\varepsilon}{1-\varepsilon}n\sigma^2.$$

From the maximum-entropy principle [30, Chap. 11], it follows that the discrete Gaussian distribution maximizes the entropy given the average energy and given the same support over a lattice. The following lemma further shows that if the flatness factor is small, the entropy rate of a discrete Gaussian $D_{\Lambda,\sigma,\mathbf{c}}$ is almost equal to the differential entropy of a continuous Gaussian of variance σ^2, minus $\frac{1}{n}\log V(\Lambda)$, that of a uniform distribution over the fundamental region of Λ.

Lemma 7 (Entropy of discrete Gaussian)**.** *Let $\mathsf{L} \sim D_{\Lambda,\sigma,\mathbf{c}}$. If $\varepsilon \triangleq \epsilon_\Lambda(\sigma/2) < 1$, then the entropy rate of L satisfies*

$$\left|\frac{1}{n}\mathbb{H}(\mathsf{L}) - \log(\sqrt{2\pi e}\sigma) - \frac{1}{n}\log V(\Lambda)\right| \le \varepsilon',$$

where $\varepsilon' = -\frac{\log(1-\varepsilon)}{n} + \frac{\pi\varepsilon}{1-\varepsilon}$.

Proof: By using the identity $f_{\sigma,\mathbf{c}}(\lambda) = \frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{\|\lambda-\mathbf{c}\|^2}{2\sigma^2}}$, we obtain:

$$\begin{aligned}\frac{1}{n}\mathbb{H}(\mathsf{L}) &= -\frac{1}{n}\sum_{\lambda\in\Lambda}\frac{f_{\sigma,\mathbf{c}}(\lambda)}{f_{\sigma,\mathbf{c}}(\Lambda)}\log\left(\frac{f_{\sigma,\mathbf{c}}(\lambda)}{f_{\sigma,\mathbf{c}}(\Lambda)}\right)\\ &= \frac{1}{n}\log\left((\sqrt{2\pi}\sigma)^n f_{\sigma,\mathbf{c}}(\Lambda)\right) + \sum_{\lambda\in\Lambda}\frac{f_{\sigma,\mathbf{c}}(\lambda)}{f_{\sigma,\mathbf{c}}(\Lambda)}\frac{\|\lambda-\mathbf{c}\|^2}{2n\sigma^2}\\ &= \frac{1}{n}\log\left((\sqrt{2\pi}\sigma)^n f_{\sigma,\mathbf{c}}(\Lambda)\right) + \frac{1}{2n\sigma^2}\mathbb{E}\left[\|\mathsf{L}-\mathbf{c}\|^2\right].\end{aligned}$$

Due to the definition of the flatness factor, we have

$$f_{\sigma,\mathbf{c}}(\Lambda) \in \left[\frac{1-\epsilon_\Lambda(\sigma)}{V(\Lambda)}, \frac{1+\epsilon_\Lambda(\sigma)}{V(\Lambda)}\right].$$

Moreover, Lemma 6 implies

$$\frac{1}{2n\sigma^2}\mathbb{E}\left[\|\mathsf{L}-\mathbf{c}\|^2\right] \in \left[\frac{1}{2} - \frac{\pi\varepsilon}{1-\varepsilon}, \frac{1}{2} + \frac{\pi\varepsilon}{1-\varepsilon}\right].$$

Since $\epsilon_\Lambda(\sigma) < \epsilon_\Lambda(\sigma/2) = \varepsilon$, we have

$$\begin{aligned}&\left|\frac{1}{n}\mathbb{H}(\mathsf{L}) - \log(\sqrt{2\pi e}\sigma) - \frac{1}{n}\log V(\Lambda)\right|\\ &\qquad < \frac{1}{n}\max\left\{\log(1+\varepsilon), -\log(1-\varepsilon)\right\} + \frac{\pi\varepsilon}{1-\varepsilon}.\end{aligned}$$

The proof is completed. □

The following lemma by Regev (adapted from [31, Claim 3.9]) shows that if the flatness factor is small, the sum of a discrete Gaussian and a continuous Gaussian is very close to a continuous Gaussian.

Lemma 8. *Let $\mathbf{c} \in \mathbb{R}^n$ be any vector, and $\sigma_0, \sigma > 0$. Consider the continuous distribution g on $\mathbb{R}^n$ obtained by adding a continuous Gaussian of variance σ^2 to a discrete Gaussian $D_{\Lambda+\mathbf{c},\sigma_0}$:*

$$g(\mathbf{x}) = \frac{1}{f_{\sigma_0}(\Lambda+\mathbf{c})}\sum_{\mathbf{t}\in\Lambda+\mathbf{c}} f_{\sigma_0}(\mathbf{t})f_\sigma(\mathbf{x}-\mathbf{t}).$$

If $\varepsilon \triangleq \epsilon_\Lambda\left(\frac{\sigma_0\sigma}{\sqrt{\sigma_0^2+\sigma^2}}\right) < \frac{1}{2}$, then $\frac{g(\mathbf{x})}{f_{\sqrt{\sigma_0^2+\sigma^2}}(\mathbf{x})}$ is uniformly close to 1:

$$\forall \mathbf{x} \in \mathbb{R}^n, \quad \left|\frac{g(\mathbf{x})}{f_{\sqrt{\sigma_0^2+\sigma^2}}(\mathbf{x})} - 1\right| \le 4\varepsilon. \tag{15}$$

In particular, the distribution $g(\mathbf{x})$ is close to the continuous Gaussian density $f_{\sqrt{\sigma_0^2+\sigma^2}}$ in L^1 distance:

$$\mathbb{V}\left(g, f_{\sqrt{\sigma_0^2+\sigma^2}}\right) \le 4\varepsilon.$$

IV. Mod-Λ Gaussian Wiretap Channel

Before considering the Gaussian wiretap channel, we will tackle a simpler model where a modulo lattice operation is performed at both the legitimate receiver's and eavesdropper's end. That is, both the legitimate channel and the eavesdropper's channel are mod-Λ channels. The mod-Λ channel is more tractable and captures the essence of the technique based on the flatness factor.

A. Channel Model

Let $\Lambda_s \subset \Lambda_e \subset \Lambda_b$ be a nested chain of n-dimensional lattices in $\mathbb{R}^n$ such that

$$\frac{1}{n}\log|\Lambda_b/\Lambda_e| = R, \quad \frac{1}{n}\log|\Lambda_e/\Lambda_s| = R'.$$

We consider the mod-Λ_s wiretap channel depicted in Figure 2. The input X^n belongs to the Voronoi region $\mathcal{V}(\Lambda_s)$ (i.e., Λ_s is the shaping lattice), while the outputs Y^n and Z^n at Bob and Eve's end respectively are given by

$$\begin{cases}\mathsf{Y}^n = [\mathsf{X}^n + \mathsf{W}_b^n] \bmod \Lambda_s,\\ \mathsf{Z}^n = [\mathsf{X}^n + \mathsf{W}_e^n] \bmod \Lambda_s,\end{cases} \tag{16}$$

where W_b^n, W_e^n are n-dimensional Gaussian vectors with zero mean and variance σ_b^2, σ_e^2 respectively.

As in the classical Gaussian channel, the transmitted codebook $\mathcal{C}$ must satisfy the average power constraint (1). We denote this wiretap channel by $W(\Lambda_s, \sigma_b, \sigma_e, P)$. Let $\mathsf{SNR}_b =$

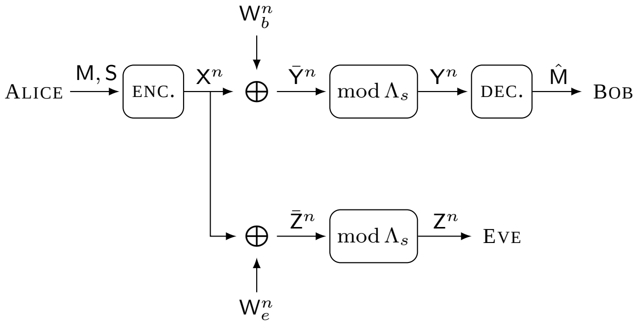


Fig. 2. The mod-Λ_s Gaussian wiretap channel.

P/σ_b^2 and $\mathsf{SNR}_e = P/\sigma_e^2$ be the signal-to-noise ratios (SNR) of Bob and Eve, respectively.

Remark 7. As was shown in [14], the capacity of a mod-Λ channel (without MMSE filtering)[3] with noise variance σ^2 is achieved by the uniform distribution on $\mathcal{V}(\Lambda)$ and is given by

$$C(\Lambda,\sigma^2)=\frac{1}{n}\left(\log(V(\Lambda))-h(\Lambda,\sigma^2)\right), \tag{17}$$

where $h(\Lambda,\sigma^2)$ is the differential entropy of the Λ-aliased noise $\bar{\mathsf{W}}^n = [\mathsf{W}^n] \bmod \Lambda$. Intuitively, the shaping lattice Λ_s must have a big flatness factor for Bob, otherwise $\bar{\mathsf{W}}^n$ will tend to a uniform distribution such that the capacity is small.

However, to the best of our knowledge, determining the secrecy capacity of the mod-Λ *wiretap* channel (16) is still an open problem. Corollary 2 in [33] provides the lower bound

$$C_s \geq C(\Lambda_s,\sigma_b^2)-C(\Lambda_s,\sigma_e^2).$$

B. Nested Lattice Codes for Binning

Consider a message set $\mathcal{M}_n = \{1,\ldots,e^{nR}\}$, and a one-to-one function $f : \mathcal{M}_n \to \Lambda_b/\Lambda_e$ which associates each message $m \in \mathcal{M}_n$ to a coset leader $\lambda_m \in \Lambda_b \cap \mathcal{V}(\Lambda_e)$. Note that we make no *a priori* assumption on the distribution of m. In order to encode the message m, Alice selects a random lattice point $\lambda \in \Lambda_e \cap \mathcal{V}(\Lambda_s)$ according to the discrete uniform distribution $p_{\mathsf{L}}(\lambda) = \frac{1}{e^{nR'}}$ and transmits $\mathsf{X}^n = \lambda + \lambda_m$. For $\bar{\lambda} \in \Lambda_e/\Lambda_s$, define

$$\begin{aligned}\mathcal{R}(\bar{\lambda}) &= \left(\mathcal{V}(\Lambda_e)+\bar{\lambda}\right) \bmod \Lambda_s\\ &= \sum_{\lambda_s\in\Lambda_s}\left(\mathcal{V}(\Lambda_e)+\bar{\lambda}+\lambda_s\right)\cap\mathcal{V}(\Lambda_s).\end{aligned}$$

The $\mathcal{R}(\bar{\lambda})$'s are fundamental regions of Λ_e and

$$\bigcup_{\bar{\lambda}\in\Lambda_s/\Lambda_e}\mathcal{R}(\bar{\lambda})=\mathcal{V}(\Lambda_s). \tag{18}$$

Figure 3 illustrates this relation by an example where $\Lambda_e = A_2$ and $\Lambda_s = 3A_2$.

To satisfy the power constraint, we choose a shaping lattice whose second moment per dimension $\sigma^2(\Lambda_s^{(n)}) = P$. Under the continuous approximation for large constellations (which could further be made precise by applying a dither), the transmission power will be equal to P.

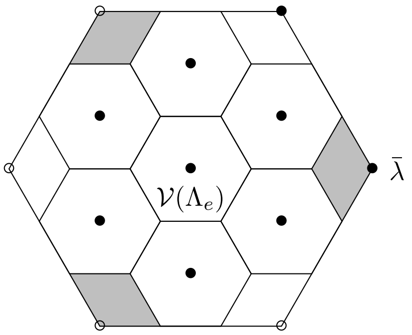


Fig. 3. The grey area represents the region $\mathcal{R}(\bar{\lambda})$ defined in (18) for the lattice pair $\Lambda_e = A_2$, $\Lambda_s = 3A_2$, with $\bar{\lambda} = (3,0)$.

C. A Sufficient Condition for Strong Secrecy

We now apply the continuous version of Csiszàr's Lemma (Lemma 1) to derive an upper bound on the amount of leaked information on the mod-Λ_s wiretap channel (16). Note that even though we consider a mod-Λ_s channel, the secrecy condition is given in terms of the flatness factor of the lattice Λ_e.

Theorem 3. *Suppose that the flatness factor of Λ_e is $\varepsilon_n \triangleq \epsilon_{\Lambda_e}(\sigma_e)$ on the eavesdropper's channel. Then*

$$\mathbb{I}(\mathsf{M};\mathsf{Z}^n) \leq 2\varepsilon_n nR - 2\varepsilon_n\log(2\varepsilon_n). \tag{19}$$

Proof: Let $\bar{\mathsf{Z}}^n = \mathsf{X}^n + \mathsf{W}_e^n$. We have, for any message m:

$$\begin{aligned}p_{\bar{\mathsf{Z}}^n|\mathsf{M}=m}(\mathbf{z}) &= \sum_{\lambda\in\Lambda_e\cap\mathcal{V}(\Lambda_s)} p_{\mathsf{L}}(\lambda)\cdot p_{\bar{\mathsf{Z}}^n|\mathsf{X}^n}(\mathbf{z}|\lambda+\lambda_m)\\ &= \frac{1}{e^{nR'}}\sum_{\lambda\in\Lambda_e\cap\mathcal{V}(\Lambda_s)} f_{\sigma_e,\lambda_m+\lambda}(\mathbf{z}).\end{aligned}$$

The output distribution of Eve's channel conditioned on m having been sent is then given by

$$\begin{aligned}p_{\mathsf{Z}^n|\mathsf{M}=m}(\mathbf{z}) &= p_{(\bar{\mathsf{Z}}^n \bmod \Lambda_s)|\mathsf{M}=m}(\mathbf{z})\\ &= \frac{1}{e^{nR'}}\sum_{\lambda\in\Lambda_e}\mathbb{1}_{\mathcal{V}(\Lambda_s)}(\mathbf{z})\cdot f_{\sigma_e,\lambda_m+\lambda}(\mathbf{z})\\ &= \frac{1}{e^{nR'}}\sum_{\bar{\lambda}\in\Lambda_e/\Lambda_s}\sum_{\lambda\in\Lambda_e}\mathbb{1}_{\mathcal{R}(\bar{\lambda})}(\mathbf{z})\cdot f_{\sigma_e,\lambda_m+\lambda}(\mathbf{z})\\ &= \frac{1}{e^{nR'}}\sum_{\bar{\lambda}\in\Lambda_e/\Lambda_s}\sum_{\lambda\in\Lambda_e}\mathbb{1}_{\mathcal{R}(\bar{\lambda})}(\mathbf{z})\cdot f_{\sigma_e,\lambda_m}(\mathbf{z}-\lambda)\\ &= \frac{1}{e^{nR'}}\sum_{\bar{\lambda}\in\Lambda_e/\Lambda_s}\bar{f}_{\bar{\lambda}}(\mathbf{z}),\end{aligned}$$

where $\bar{f}_{\bar{\lambda}}(\mathbf{z}) = \sum_{\lambda\in\Lambda_e}\mathbb{1}_{\mathcal{R}(\bar{\lambda})}(\mathbf{z})\cdot f_{\sigma_e,\lambda_m}(\mathbf{z}-\lambda)$ is the density function of a continuous Gaussian with variance σ_e^2 and center λ_m reduced modulo the fundamental region $\mathcal{R}(\bar{\lambda})$. From Proposition 3, we have that $\mathbb{V}(\bar{f}_{\bar{\lambda}}, U_{\mathcal{R}(\bar{\lambda})}) \leq \epsilon_{\Lambda_e}(\sigma_e)$

[3] It is known that if an MMSE filter is added before the mod-Λ operation, there exists a sequence of lattices approaching the capacity of the AWGN channel [15, 32]. However, MMSE filtering is not considered in this section.

for all $\bar{\lambda} \in \Lambda_e/\Lambda_s$. From the decomposition $U_{\mathcal{V}(\Lambda_s)}(\mathbf{z}) = \frac{1}{e^{nR'}} \sum_{\bar{\lambda} \in \Lambda_e/\Lambda_s} U_{\mathcal{R}(\bar{\lambda})}(\mathbf{z})$, we obtain

$$\begin{aligned}&\mathbb{V}(p_{\mathsf{Z}^n|\mathsf{M}=m}, U_{\mathcal{V}(\Lambda_s)}) \\ &\qquad \leq \frac{1}{e^{nR'}} \sum_{\bar{\lambda} \in \Lambda_e/\Lambda_s} \int_{R(\bar{\lambda})} \left| \bar{f}_{\bar{\lambda}}(\mathbf{z}) - U_{\mathcal{R}(\bar{\lambda})}(\mathbf{z}) \right| d\mathbf{z} \\ &\qquad \leq \epsilon_{\Lambda_e}(\sigma_e).\end{aligned}$$

Recalling the definition of d_{av} in Lemma 1, defining $q_{\mathsf{Z}}(\mathbf{z}) = U_{\mathcal{V}(\Lambda_s)}(\mathbf{z})$, and using the inequality (6), we find that $d_{\text{av}} \leq 2\epsilon_{\Lambda_e^{(n)}}(\sigma_e)$. Then the mutual information can be estimated using Lemma 2. □

From Theorem 3, we obtain a sufficient condition for a sequence of nested lattice wiretap codes to achieve strong secrecy.

Corollary 1. *For any sequence of lattices $\Lambda_e^{(n)}$ such that $\epsilon_{\Lambda_e^{(n)}}(\sigma_e) = o\left(\frac{1}{n}\right)$ as $n \to \infty$, we have $\mathbb{I}(\mathsf{M}; \mathsf{Z}^n) \to 0$.*

In fact, Theorem 2 guarantees the existence of mod-p lattices $\Lambda_e^{(n)}$ whose flatness factor is exponentially small. Therefore, if Eve's generalized SNR $\gamma_{\Lambda_e}(\sigma_e)$ is smaller than 1, then strong secrecy can be achieved by such lattice codes, and in that setup the mutual information will vanish exponentially fast. We say that such lattices achieve σ_e-secrecy.

Now, we introduce the notion of secrecy-good lattices. Roughly speaking, a lattice is good for secrecy if its flatness factor is small. Although $\epsilon_{\Lambda_e^{(n)}}(\sigma_e) = o\left(\frac{1}{n}\right)$ is sufficient to achieve strong secrecy, it is desired in practice that the information leakage is exponentially small. Thus, we define secrecy-goodness as follows:

Definition 9 (Secrecy-good). *A sequence of lattices $\Lambda^{(n)}$ is* secrecy-good *if*

$$\epsilon_{\Lambda^{(n)}}(\sigma) \leq 2^{-\Omega(n)}, \quad \forall \gamma_{\Lambda^{(n)}}(\sigma) < 1. \tag{20}$$

This property is invariant with respect to scaling, i.e., when both σ and Λ are scaled accordingly (cf. Remark 5). It is slightly more general than (12) of Theorem 2. The purpose is to accommodate the lattices whose theta series are close to, but not exactly below the Minkowski-Hlawka bound.

Alternatively, we may employ the smoothing parameter to state the secrecy goodness of lattices. From (12), we have that on average the smoothing parameter of mod-p lattices is given by

$$\bar{\eta}_\varepsilon = \left(\frac{(1+\delta)V(\Lambda)}{\varepsilon} \right)^{1/n} \to V(\Lambda)^{1/n}$$

for any fixed ε. So, equivalently a sequence of lattices $\Lambda^{(n)}$ is secrecy-good if the smoothing parameter is smaller than or equal to $\bar{\eta}_\varepsilon$. In other words, $\Lambda^{(n)}$ has a threshold of noise standard deviation σ_e smaller than or equal to $\frac{V(\Lambda)^{1/n}}{\sqrt{2\pi}}$ beyond which the amount of information leakage vanishes.

D. Existence of Good Wiretap Codes from Nested Lattices

A priori, the secrecy-goodness property established in the previous subsection may come at the expense of reliability for the legitimate receiver. We will show that this is not the case, i.e., that there exists a sequence of nested lattices which guarantee both strong secrecy rates and reliability:

Proposition 4. *Given $R, R' > 0$, there exists a sequence of nested lattices $\Lambda_s^{(n)} \subset \Lambda_e^{(n)} \subset \Lambda_b^{(n)}$ whose nesting ratios satisfy*

$$R'_n = \frac{1}{n} \log \frac{V(\Lambda_s)}{V(\Lambda_e)} \to R', \quad R_n = \frac{1}{n} \log \frac{V(\Lambda_e)}{V(\Lambda_b)} \to R$$

when $n \to \infty$, and such that

- *$\Lambda_s^{(n)}$ is quantization and AWGN-good,*
- *$\Lambda_e^{(n)}$ is secrecy-good,*
- *$\Lambda_b^{(n)}$ is AWGN-good.*

The proof of Proposition 4 can be found in Appendix III and follows the approach of [34]. The main novelty is the addition of the secrecy-goodness property, which requires checking that the corresponding condition is compatible with the ones introduced in [34].

Theorem 4. *Let $\sigma_e^2 > e \cdot \sigma_b^2$. Then as $n \to \infty$, all strong secrecy rates R satisfying*

$$R < \frac{1}{2} \log \frac{\sigma_e^2}{\sigma_b^2} - \frac{1}{2}$$

are achievable using nested lattice codes $\Lambda_s^{(n)} \subset \Lambda_e^{(n)} \subset \Lambda_b^{(n)}$ on the mod-$\Lambda_s^{(n)}$ wiretap channel $W(\Lambda_s, \sigma_b, \sigma_e, P)$.

Proof: Consider the binning scheme described in Section IV-B, where the nested lattices $\Lambda_s^{(n)} \subset \Lambda_e^{(n)} \subset \Lambda_b^{(n)}$ are given by Proposition 4. Since $\Lambda_b^{(n)}$ is AWGN-good, without MMSE filtering, a channel coding rate (without secrecy constraint) $R + R' < \frac{1}{2} \log \mathsf{SNR}_b$ is achievable at the legitimate receiver's end, with the error probability vanishing exponentially fast in n [15].

Since $\Lambda_e^{(n)}$ is secrecy-good, by Theorem 2 in order to have strong secrecy at the eavesdropper's end, it is sufficient for mod-p lattices to have

$$\gamma_{\Lambda_e}(\sigma_e) = \frac{V(\Lambda_s)^{\frac{2}{n}}}{(e^{nR'})^{\frac{2}{n}} 2\pi\sigma_e^2} \to \frac{P \cdot e}{e^{2R'}\sigma_e^2} < 1,$$

where $V(\Lambda_s)^{\frac{2}{n}} \to 2\pi e \sigma^2(\Lambda_s^{(n)})$ because $\Lambda_s^{(n)}$ is quantization-good and also $P = \sigma^2(\Lambda_s^{(n)})$ under the continuous approximation. The above relation implies

$$R' > \frac{1}{2} \log \mathsf{SNR}_e + \frac{1}{2}. \tag{21}$$

Consequently, all strong secrecy rates R satisfying

$$R < \frac{1}{2} \log \frac{\sigma_e^2}{\sigma_b^2} - \frac{1}{2}$$

are achievable on the wiretap channel $W(\Lambda_s, \sigma_b, \sigma_e, P)$. Note that positive rates are achievable by the proposed scheme only if $\sigma_b^2 > e \cdot \sigma_e^2$. □

For high SNR, the strong secrecy rate that can be achieved using Proposition 4 is very close to the lower bound on the secrecy capacity, to within a half nat.

Remark 8. In our strong secrecy scheme, the output distribution of each bin with respect to the eavesdropper's channel approaches the output of the uniform distribution in variational distance. That is, each bin is a *resolvability code* in the sense of Han and Verdú [35]. In [17, 18] it was shown that for discrete memoryless channels, resolvability-based random wiretap codes achieve strong secrecy; we have followed a similar approach for the Gaussian channel.

In the case when the target output distribution is capacity-achieving, a necessary condition for the bins to be resolvability codes is that the bin rate should be greater than the eavesdropper's channel capacity. Note that this is consistent with the condition (21): if Λ_s is good for quantization, the entropy of the Λ_s-aliased noise $\bar{\mathsf{W}}^n = [\mathsf{W}^n] \bmod \Lambda_s$ tends to the entropy of a white Gaussian noise with the same variance [36], and $V(\Lambda_s) \approx (2\pi e P)^{\frac{n}{2}}$, so the capacity $C(\Lambda_s, \sigma_e^2)$ of the eavesdropper's channel given by equation (17) tends to $\frac{1}{2}\log 2\pi e P - \frac{1}{2}\log 2\pi e \sigma_e^2 = \frac{1}{2}\log \mathsf{SNR}_e$.

Remark 9 (Relation to Poltyrev's setting of infinite constellations)**.** Poltyrev initiated the study of infinite constellations in the presence of Gaussian noise [28]. In this setting, although the standard channel capacity is meaningless (so he defined generalized capacity), the secrecy capacity is finite. This is because the secrecy capacity of the Gaussian wiretap channel as $P \to \infty$ converges to a finite rate $\frac{1}{2}\log(\frac{\sigma_e^2}{\sigma_b^2})$. Lattice codes can not be better than this, so it is an upper bound. Even though we considered a mod-Λ_s channel in this section, we may enlarge $\mathcal{V}(\Lambda_s)$ (i.e., increase R' while fixing R) to approach an infinite constellation. Since the upper bound (19) on the mutual information of our proposed scheme is independent of $V(\Lambda_s)$, the limit exists as $V(\Lambda_s) \to \infty$. This corresponds to the case of infinite constellations. Further, the achieved secrecy rate is only a half nat away from the upper bound.

V. Gaussian Wiretap Channel

Although the mod-Λ channel has led to considerable insights, there is no reason in real-world applications why the eavesdropper would be restricted to use the modulo operation in the front end of her receiver. In this section, we remove this restriction and solve the problem of the Gaussian wiretap channel using lattice Gaussian signaling.

A. Channel Model

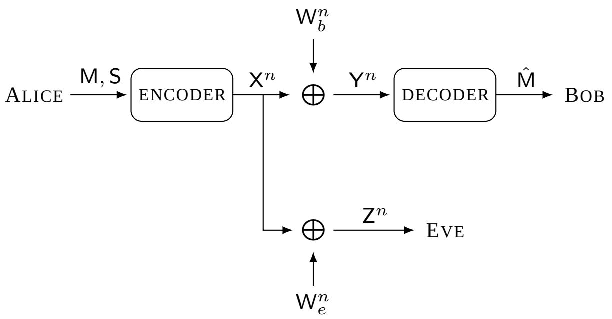


Fig. 4. The Gaussian wiretap channel.

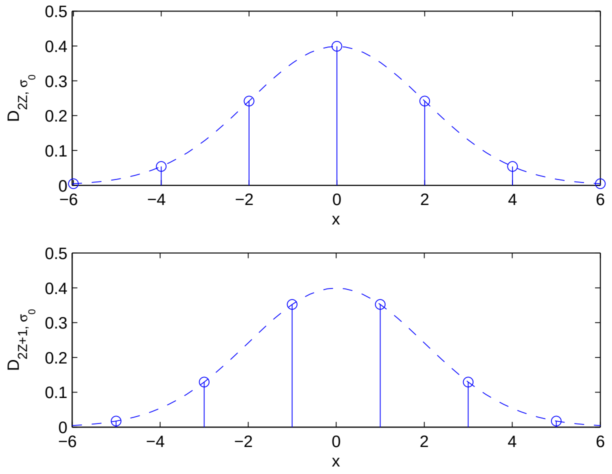


Fig. 5. Lattice Gaussian signaling (circle) over $2\mathbb{Z}$ and its coset $2\mathbb{Z}+1$ for $\sigma_0 = 2$. The profile (dashed) is the underlying continuous Gaussian distribution.

Let $\Lambda_e \subset \Lambda_b$ be n-dimensional lattices in $\mathbb{R}^n$ such that

$$\frac{1}{n}\log|\Lambda_b/\Lambda_e| = R.$$

We consider the Gaussian wiretap channel depicted in Fig. 4, whose outputs Y^n and Z^n at Bob and Eve's end respectively are given by

$$\begin{cases} \mathsf{Y}^n = \mathsf{X}^n + \mathsf{W}_b^n, \\ \mathsf{Z}^n = \mathsf{X}^n + \mathsf{W}_e^n, \end{cases} \tag{22}$$

where W_b^n, W_e^n are n-dimensional Gaussian vectors with zero mean and variance σ_b^2, σ_e^2 respectively. The transmitted codebook $\mathcal{C}$ must satisfy the average power constraint (1). We denote this wiretap channel by $W(\sigma_b, \sigma_e, P)$. Again, let $\mathsf{SNR}_b = P/\sigma_b^2$ and $\mathsf{SNR}_e = P/\sigma_e^2$.

B. Lattice Gaussian Signaling

Consider a message set $\mathcal{M}_n = \{1, \ldots, e^{nR}\}$, and a one-to-one function $\phi : \mathcal{M}_n \to \Lambda_b/\Lambda_e$ which associates each message $m \in \mathcal{M}_n$ to a coset $\bar{\lambda}_m \in \Lambda_b/\Lambda_e$. One could choose the coset leader $\lambda_m \in \Lambda_b \cap \mathcal{R}(\Lambda_e)$ for any fundamental region $\mathcal{R}(\Lambda_e)$, not necessarily the Voronoi region $\mathcal{V}(\Lambda_e)$. This is because the signal powers corresponding to different m will be nearly the same, as shown in the following. This property can result in convenient implementation of the encoder. Note again that we make no *a priori* assumption on the distribution of m.

In order to encode the message $m \in \mathcal{M}_n$, Alice samples X_m^n from $D_{\Lambda_e+\lambda_m,\sigma_0}$ (as defined in Section III-C); equivalently, Alice transmits $\lambda + \lambda_m$ where $\lambda \sim D_{\Lambda_e,\sigma_0,-\lambda_m}$. The choice of the variance σ_0^2 will be discussed later in this Section.

It is worth mentioning that the distribution $D_{\Lambda_e+\lambda_m,\sigma_0}$ is always centered at $\mathbf{0}$ for all bins. Fig. 5 illustrates the proposed lattice Gaussian signaling using an example $\Lambda_e = 2\mathbb{Z}$ for $\sigma_0 = 2$. It is clear that both $D_{2\mathbb{Z},\sigma_0}$ and $D_{2\mathbb{Z}+1,\sigma_0}$ are centered at 0, sharing the same continuous Gaussian profile. This is key for the conditional output distributions corresponding to different m to converge to the same distribution.

Lemma 6 implies that if $\epsilon_{\Lambda_e}(\sigma_0/2) < 1/2$, then

$$\left|\mathbb{E}\left[\|\mathsf{X}_m^n\|^2\right] - n\sigma_0^2\right| \le \frac{2\pi\epsilon_{\Lambda_e}(\sigma_0/2)}{1-\epsilon_{\Lambda_e}(\sigma_0/2)} n\sigma_0^2,$$

which is independent of m. Note that the overall input distribution is a mixture of the densities of X_m^n:

$$p_{\mathsf{X}^n}(\mathbf{x}) = \sum_{m=1}^{e^{nR}} p_{\mathsf{M}}(m) p_{\mathsf{X}_m^n}(\mathbf{x}). \tag{23}$$

Since the second moment in zero of a mixture of densities is the weighted sum of the second moments in zero of the individual densities, we have

$$\left|\frac{1}{n}\mathbb{E}\left[\|\mathsf{X}^n\|^2\right] - \sigma_0^2\right| \le \frac{2\pi\epsilon_{\Lambda_e}(\sigma_0/2)}{1-\epsilon_{\Lambda_e}(\sigma_0/2)} \sigma_0^2. \tag{24}$$

We choose $\sigma_0^2 = P$ in order to satisfy the average power constraint (1) asymptotically (as $\epsilon_{\Lambda_e}(\sigma_0/2) \to 0$). For convenience, let $\rho_b = \sigma_0^2/\sigma_b^2$ and $\rho_e = \sigma_0^2/\sigma_e^2$. It holds that $\rho_b \to \mathsf{SNR}_b$ and $\rho_e \to \mathsf{SNR}_e$ if $\epsilon_{\Lambda_e}(\sigma_0/2) \to 0$.

C. Achieving Strong Secrecy

We will now show that under suitable hypotheses, the conditional output distributions at Eve's end converge in variational distance to the same continuous Gaussian distribution, thereby achieving strong secrecy.

Recall that Eve's channel transition probability is given by

$$p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathbf{z}|\lambda_m + \lambda) = f_{\sigma_e,\lambda_m+\lambda}(\mathbf{z}).$$

Let $\tilde{\sigma}_e = \frac{\sigma_0\sigma_e}{\sqrt{\sigma_0^2+\sigma_e^2}}$. Lemma 8 implies that if $\epsilon_{\Lambda_e}(\tilde{\sigma}_e) < \frac{1}{2}$, then:

$$\mathbb{V}\left(p_{\mathsf{Z}^n|\mathsf{M}}(\cdot|m), f_{\sqrt{\sigma_0^2+\sigma_e^2}}\right) \le 4\epsilon_{\Lambda_e}(\tilde{\sigma}_e).$$

An upper bound on the amount of leaked information then follows directly from Lemma 2.

Theorem 5. *Suppose that the wiretap coding scheme described above is employed on the Gaussian wiretap channel (22), and let $\varepsilon_n = \epsilon_{\Lambda_e}(\tilde{\sigma}_e)$. Assume that $\varepsilon_n < \frac{1}{2}$ for all n. Then the mutual information between the confidential message and the eavesdropper's signal is bounded as follows:*

$$\mathbb{I}(\mathsf{M};\mathsf{Z}^n) \le 8\varepsilon_n nR - 8\varepsilon_n \log 8\varepsilon_n \tag{25}$$

From Theorem 5, we obtain a sufficient condition for a sequence of nested lattice wiretap codes to achieve strong secrecy:

Corollary 2. *For any sequence of lattices $\Lambda_e^{(n)}$ such that $\epsilon_{\Lambda_e^{(n)}}(\tilde{\sigma}_e) = o\left(\frac{1}{n}\right)$ as $n \to \infty$, we have $\mathbb{I}(\mathsf{M},\mathsf{Z}^n) \to 0$.*

Note that $\tilde{\sigma}_e$ is smaller than both σ_e and σ_0. The first inequality $\tilde{\sigma}_e < \sigma_e$ means that

- Because of the monotonicity of the flatness factor (Remark 3), achieving strong secrecy on the Gaussian wiretap channel is a bit more demanding than that on the mod-Λ channel;
- Yet they are equally demanding at high SNR, since $\tilde{\sigma}_e \to \sigma_e$ as $\sigma_0 \to \infty$.

The second inequality $\tilde{\sigma}_e < \sigma_0$ requires that $\epsilon_{\Lambda_e}(\sqrt{P})$ be small, which means that a minimum power P is needed (specifically, $\sqrt{P}$ should be larger than the smoothing parameter of Λ_e).

Remark 10. Note that, similarly to the mod-Λ case (Remark 8) each bin of our strong secrecy scheme may be viewed as a resolvability code, and thus the bin rate must necessarily be above Eve's channel capacity. Indeed, the bin rate can be chosen to be quite close to this optimal value: note that for ε_n in Theorem 5 to vanish, it suffices that

$$\gamma_{\Lambda_e}(\tilde{\sigma}_e) = \frac{V(\Lambda_e)^{2/n}}{2\pi\tilde{\sigma}_e^2} < 1 \tag{26}$$

for the mod-p lattices of the first part of Theorem 2. By Proposition 7, when $\varepsilon \triangleq \epsilon_{\Lambda_e}(\sigma_0/2) < 1$, the entropy rate of each bin satisfies

$$\begin{aligned} R' &\ge \log(\sqrt{2\pi e}\sigma_0) - \frac{1}{n}\log V(\Lambda_e) - \varepsilon' \\ &> \log(\sqrt{2\pi e}\sigma_0) - \frac{1}{2}\log\left(2\pi\frac{\sigma_0^2\sigma_e^2}{\sigma_0^2+\sigma_e^2}\right) - \varepsilon' \\ &= \frac{1}{2}\log\left(\frac{\sigma_0^2+\sigma_e^2}{\sigma_e^2}\right) + \frac{1}{2} - \varepsilon' \\ &= \frac{1}{2}\log(1+\rho_e) + \frac{1}{2} - \varepsilon'. \end{aligned}$$

where ε' is defined in Proposition 7. Since $P \to \sigma_0^2$ as $\varepsilon \to 0$ (by (24)), we have $\rho_e \to \mathsf{SNR}_e$. Also, $\varepsilon' \to 0$ as $\varepsilon \to 0$. To make $\varepsilon \to 0$, we only need an extra sufficient condition $\gamma_{\Lambda_e}(\sigma_0/2) < 1$ for the mod-p lattices of Theorem 2.

D. Achieving Reliability

Now we show Bob can reliably decode the confidential message by using MMSE lattice decoding. Consider the decoding scheme for Bob where he first decodes to the fine lattice Λ_b, then applies the mod-Λ_e operation to recover the confidential message. We note that the distribution of Alice's signal can be approximated by D_{Λ_b,σ_0}, when the confidential message is uniformly distributed. More precisely, since Alice transmits $\mathbf{x} \sim D_{\Lambda_e+\lambda_m,\sigma_0}$, by Lemma 5, the density p_{X^n} of $\mathbf{x}$ is close to the discrete Gaussian distribution over Λ_b, if $\lambda_m \in \Lambda_b/\Lambda_e$ is uniformly distributed. In fact, we have $\mathbb{V}(p_{\mathsf{X}^n}, D_{\Lambda_b,\sigma_0}) \le \frac{2\varepsilon}{1-\varepsilon}$ when $\varepsilon \triangleq \epsilon_{\Lambda_e}(\sigma_0) < \frac{1}{2}$.

We will derive the maximum-a-posteriori (MAP) decoding rule for decoding to Λ_b, assuming a discrete Gaussian distribution D_{Λ_b,σ_0} over Λ_b. Since the lattice points are not equally probable a priori in the lattice Gaussian signaling, MAP decoding is not the same as standard maximum-likelihood (ML) decoding.

Proposition 5 (Equivalence between MAP decoding and MMSE lattice decoding)**.** *Let $\mathbf{x} \sim D_{\Lambda_b,\sigma_0}$ be the input signaling of an AWGN channel where the noise variance is σ_b^2. Then MAP decoding is equivalent to Euclidean lattice decoding of Λ_b using a renormalized metric that is asymptotically close to the MMSE metric.*

Proof: Bob receives $\mathbf{y} = \mathbf{x} + \mathbf{w}_b$. Thus the MAP decoding metric is given by

$$\begin{aligned}\mathbb{P}(\mathbf{x}|\mathbf{y}) = \frac{\mathbb{P}(\mathbf{x},\mathbf{y})}{\mathbb{P}(\mathbf{y})} &\propto \mathbb{P}(\mathbf{y}|\mathbf{x})\mathbb{P}(\mathbf{x}) \\ &\propto \exp\left(-\frac{\|\mathbf{y}-\mathbf{x}\|^2}{2\sigma_b^2} - \frac{\|\mathbf{x}\|^2}{2\sigma_0^2}\right) \\ &\propto \exp\left(-\frac{1}{2}\left(\frac{\sigma_0^2+\sigma_b^2}{\sigma_0^2\sigma_b^2}\left\|\frac{\sigma_0^2}{\sigma_0^2+\sigma_b^2}\mathbf{y}-\mathbf{x}\right\|^2\right)\right).\end{aligned}$$

Therefore,

$$\begin{aligned}\arg\max_{\mathbf{x}\in\Lambda_b}\mathbb{P}(\mathbf{x}|\mathbf{y}) &= \arg\min_{\mathbf{x}\in\Lambda_b}\left\|\frac{\sigma_0^2}{\sigma_0^2+\sigma_b^2}\mathbf{y}-\mathbf{x}\right\|^2 \\ &= \arg\min_{\mathbf{x}\in\Lambda_b}\|\alpha\mathbf{y}-\mathbf{x}\|^2 \qquad (27)\end{aligned}$$

where $\alpha = \frac{\sigma_0^2}{\sigma_0^2+\sigma_b^2}$ is known, thanks to (24), to be asymptotically close to the MMSE coefficient $\frac{P}{P+\sigma_b^2}$. □

Next we prove Bob's reliability for any secrecy rate close to the secrecy capacity. We use the α-renormalized decoding metric (27), even if the confidential message is not necessarily uniformly distributed. In fact, the following proofs hold for any fixed message index m. Also note that no dither is required to achieve reliability. Indeed, as we will see, Regev's regularity lemma (Lemma 8) makes the dither unnecessary. This is because the equivalent noise will be asymptotically Gaussian.

Suppose Alice transmits message m, and Bob receives $\mathbf{y} = \mathbf{x} + \mathbf{w}_b = \lambda + \lambda_m + \mathbf{w}_b$ (with $\lambda \sim D_{\Lambda_e,\sigma_0,-\lambda_m}$). From Proposition 5, Bob computes

$$\hat{\lambda}_m = [Q_{\Lambda_b}(\alpha\mathbf{y})] \bmod \Lambda_e.$$

Recall the following properties of the $\bmod \Lambda$ and quantization operations. For all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, we have

$$[[\mathbf{a}] \bmod \Lambda_e + \mathbf{b}] \bmod \Lambda_e = [\mathbf{a}+\mathbf{b}] \bmod \Lambda_e \qquad (28)$$

$$[Q_{\Lambda_b}(\mathbf{a})] \bmod \Lambda_e = [Q_{\Lambda_b}([\mathbf{a}] \bmod \Lambda_e)] \bmod \Lambda_e. \qquad (29)$$

Using these properties, the output of Bob's decoder can be rewritten as

$$\begin{aligned}\hat{\lambda}_m &= [Q_{\Lambda_b}(\mathbf{x} + (\alpha-1)\,\mathbf{x} + \alpha\mathbf{w}_b)] \bmod \Lambda_e \\ &= [Q_{\Lambda_b}([\mathbf{x} + (\alpha-1)\,\mathbf{x} + \alpha\mathbf{w}_b] \bmod \Lambda_e)] \bmod \Lambda_e.\end{aligned}$$

Observe that since $\lambda \in \Lambda_e$, we have

$$\begin{aligned}&[\mathbf{x} + (\alpha-1)\,\mathbf{x} + \alpha\mathbf{w}_b] \bmod \Lambda_e \\ =&[\lambda_m + (\alpha-1)\,\mathbf{x} + \alpha\mathbf{w}_b] \bmod \Lambda_e \\ =&[\lambda_m + \tilde{\mathbf{w}}_b(m)] \bmod \Lambda_e\end{aligned}$$

where we have defined the equivalent noise

$$\tilde{\mathbf{w}}_b(m) = (\alpha-1)\,\mathbf{x} + \alpha\mathbf{w}_b.$$

Therefore

$$\hat{\lambda}_m = [Q_{\Lambda_b}(\lambda_m + \tilde{\mathbf{w}}_b(m))] \bmod \Lambda_e.$$

Let $p_{\tilde{\mathsf{W}}_b^n(m)}$ be the density of the equivalent noise $\tilde{\mathbf{w}}_b(m)$. Since $\mathbf{x} \sim D_{\Lambda_e+\lambda_m,\sigma_0}$ and $\mathbf{w}_b$ is Gaussian, Lemma 8 implies that for *any fixed* m, and randomizing over λ, $p_{\tilde{\mathsf{W}}_b^n(m)}$ is very close to a continuous Gaussian distribution. More precisely, applying Lemma 8 with standard deviations $(\alpha-1)\sigma_0$ and $\alpha\sigma_b$, and defining $\tilde{\sigma}_b = \sqrt{(\alpha-1)^2\sigma_0^2 + \alpha^2\sigma_b^2} = \frac{\sigma_0\sigma_b}{\sqrt{\sigma_0^2+\sigma_b^2}}$, we have

$$\left|p_{\tilde{\mathsf{W}}_b^n(m)}(\mathbf{w}) - f_{\tilde{\sigma}_b}(\mathbf{w})\right| \leq 4\varepsilon'' f_{\tilde{\sigma}_b}(\mathbf{w}) \quad \forall \mathbf{w} \in \mathbb{R}^n, \qquad (30)$$

assuming that (recall $\rho_b = \sigma_0^2/\sigma_b^2$)

$$\varepsilon'' \triangleq \epsilon_{(1-\alpha)\Lambda_e}\left(\frac{(1-\alpha)\sigma_0}{\sqrt{1+1/\rho_b}}\right) = \epsilon_{\Lambda_e}\left(\frac{\sigma_0}{\sqrt{1+1/\rho_b}}\right) < \frac{1}{2}.$$

Thus, if $\varepsilon'' \to 0$, the equivalent noise is essentially statistically independent from m, in the sense that it is very close to the distribution $f_{\tilde{\sigma}_b}(\mathbf{w})$ that does not involve m at all.

Theorem 6. *Suppose* $\mathsf{SNR}_b > 4e - 1$, $\frac{1+\mathsf{SNR}_b}{1+\mathsf{SNR}_e} > e$, *and* $\mathsf{SNR}_b \cdot \mathsf{SNR}_e > 1$. *Then if* $\Lambda_b^{(n)}$ *is a sequence of AWGN-good lattices, and* $\Lambda_e^{(n)}$ *is a sequence of secrecy-good lattices, any strong secrecy rate* R *satisfying*

$$R < \frac{1}{2}\log(1+\mathsf{SNR}_b) - \frac{1}{2}\log(1+\mathsf{SNR}_e) - \frac{1}{2} \qquad (31)$$

is achievable on the Gaussian wiretap channel $W(\sigma_b, \sigma_e, P)$ *using the discrete Gaussian signaling and MMSE-renormalized Euclidean lattice decoding.*

Proof: The decoding error probability $P_e(m)$ corresponding to the message m is bounded from above as

$$\begin{aligned}P_e(m) \leq \mathbb{P}\left\{Q_{\Lambda_b}(\lambda_m + \tilde{\mathbf{w}}_b(m)) \neq \lambda_m\right\}& \\ = \mathbb{P}\left\{\tilde{\mathbf{w}}_b(m) \notin \mathcal{V}(\Lambda_b)\right\}&.\end{aligned}$$

Since in particular

$$p_{\tilde{\mathsf{W}}_b^n(m)}(\mathbf{w}) < (1+4\varepsilon'') f_{\tilde{\sigma}_b}(\mathbf{w}) \quad \forall \mathbf{w} \in \mathbb{R}^n,$$

we find that

$$\mathbb{P}\left\{\tilde{\mathbf{w}}_b(m) \notin \mathcal{V}(\Lambda_b)\right\} \leq (1+4\varepsilon'') \cdot \mathbb{P}\left\{\hat{\mathbf{w}}_b \notin \mathcal{V}(\Lambda_b)\right\}$$

where $\hat{\mathbf{w}}_b$ is i.i.d. Gaussian with variance $\tilde{\sigma}_b^2$. Note that while the equivalent noise $\tilde{\mathbf{w}}_b(m)$ in general depends on m, the resulting bound on the error probability is independent of m.

From AWGN-goodness of Λ_b, it follows that the decoding error probability P_e tends to 0 exponentially fast if ε'' is bounded by a constant and if

$$\gamma_{\Lambda_b}(\tilde{\sigma}_b) = \frac{V(\Lambda_b)^{2/n}}{2\pi\tilde{\sigma}_b^2} > e. \qquad (32)$$

On the other hand, since Λ_e is secrecy-good, Theorem 5 implies that a sufficient condition for the mod-p lattices of Theorem 2 to achieve strong secrecy is

$$\gamma_{\Lambda_e}(\tilde{\sigma}_e) = \frac{V(\Lambda_e)^{2/n}}{2\pi\tilde{\sigma}_e^2} < 1. \qquad (33)$$

Combining (32) and (33), we have that strong secrecy rates R satisfying

$$R = \frac{1}{n}\log\frac{V(\Lambda_e)}{V(\Lambda_b)} < \frac{1}{2}\log\left(\frac{1+\rho_b}{1+\rho_e}\right) - \frac{1}{2} \qquad (34)$$

are achievable.

Two extra conditions on the flatness factors are required. First, to make $\rho_b \to \mathsf{SNR}_b$ and $\rho_e \to \mathsf{SNR}_e$, it suffices that $\epsilon_{\Lambda_e}(\sigma_0/2) \to 0$ (by (24)). This condition can be satisfied by mod-p lattices if

$$\gamma_{\Lambda_e}\left(\frac{\sigma_0}{2}\right) = \frac{V(\Lambda_e)^{2/n}}{2\pi\frac{\sigma_0^2}{4}} < 1,$$

which together with (32) limits the secrecy rate to

$$R < \frac{1}{2}\log\left(\frac{1+\rho_b}{4}\right) - \frac{1}{2}. \tag{35}$$

The second condition $\epsilon_{\Lambda_e}\left(\frac{\sigma_0}{\sqrt{1+1/\rho_b}}\right) \to 0$ for the equivalent noise to be asymptotically Gaussian (by (30)) can be satisfied by mod-p lattices if

$$\gamma_{\Lambda_e}\left(\frac{\sigma_0}{\sqrt{1+1/\rho_b}}\right) = \frac{V(\Lambda_e)^{2/n}}{2\pi\frac{\sigma_0^2}{1+1/\rho_b}} < 1,$$

which together with (32) limits the secrecy rate to

$$R < \frac{1}{2}\log\rho_b - \frac{1}{2}. \tag{36}$$

Now, combining (34)-(36) and considering a positive secrecy rate, we have

$$R < \frac{1}{2}\log\left(\min\left\{\frac{1+\mathsf{SNR}_b}{1+\mathsf{SNR}_e}, \mathsf{SNR}_b\right\}\right) - \frac{1}{2} \tag{37}$$

when $\mathsf{SNR}_b > 4e - 1$ and $\frac{1+\mathsf{SNR}_b}{1+\mathsf{SNR}_e} > e$. Note that condition (35) has been absorbed in (37). Further, when $\mathsf{SNR}_b \cdot \mathsf{SNR}_e > 1$, the first term is smaller. Therefore, the theorem is proven.

□

Remark 11. It can be checked that, in our framework, conventional (non-renormalized) minimum-distance lattice decoding can only achieve strong secrecy rate up to

$$R < \frac{1}{2}\log\left(\mathsf{SNR}_b\right) - \frac{1}{2}\log\left(1+\mathsf{SNR}_e\right) - \frac{1}{2}.$$

This is because it requires

$$\gamma_{\Lambda_b}(\sigma_b) = \frac{V(\Lambda_b)^{2/n}}{2\pi\sigma_b^2} > e$$

rather than (32). Therefore, MAP decoding or MMSE estimation allows to gain a constant 1 within the logarithm of the first term.

Remark 12. The existence of good wiretap codes for the Gaussian channel follows from Proposition 4. In fact, this case is less demanding than the mod-Λ_s channel there since no shaping lattice is needed. We only need a sequence of nested lattices $\Lambda_e^{(n)} \subset \Lambda_b^{(n)}$ where $\Lambda_e^{(n)}$ is secrecy-good (with respect to $\tilde{\sigma}_e$ rather than σ_e) and $\Lambda_b^{(n)}$ is AWGN-good.

VI. Discussion

In this paper, we have studied semantic security over the Gaussian wiretap channel using lattice codes. The flatness factor serves as a new lattice parameter to measure information leakage in this setting. It can tell whether a particular lattice is good or not for secrecy coding, and consequently provides a design criterion of wiretap lattice codes. While we have proved the existence of secrecy-good mod-p lattices, the explicit construction of practical secrecy-good lattices warrants an investigation. Further work along the line of secrecy gain [10] may provide some secrecy-good unimodular lattices.

The half-nat gap to the secrecy capacity is intriguing. It would be interesting to find out what happens in between, and to further explore the relation between various lattice parameters.

Acknowledgments

The authors would like to thank Matthieu Bloch and Guillaume Hanrot for helpful discussions.

Appendix I
Proof of Csiszár's Lemma for Continuous Channels

Proof: Note that in spite of the ambiguous notation, here p_{Z} and $p_{\mathsf{Z}|\mathsf{M}=m}$ are densities on $\mathbb{R}^n$, while p_{M} and $p_{\mathsf{M}|\mathsf{Z}=\mathbf{z}}$ are probability mass functions on $\mathcal{M}_n$. We have

$$\begin{aligned}
d_{\text{av}} &= \sum_{m\in\mathcal{M}_n} p_{\mathsf{M}}(m)\int_{\mathbb{R}^n}\left|p_{\mathsf{Z}|\mathsf{M}=m}(z) - p_{\mathsf{Z}}(z)\right| dz \\
&= \sum_{m\in\mathcal{M}_n}\int_{\mathbb{R}^n}\left|p_{\mathsf{M}|\mathsf{Z}=\mathbf{z}}(m)p_{\mathsf{Z}}(\mathbf{z}) - p_{\mathsf{M}}(m)p_{\mathsf{Z}}(\mathbf{z})\right| d\mathbf{z} \\
&= \int_{\mathbb{R}^n}\sum_{m\in\mathcal{M}_n}\left|p_{\mathsf{M}|\mathsf{Z}=\mathbf{z}}(m) - p_{\mathsf{M}}(m)\right| p_{\mathsf{Z}}(\mathbf{z})d\mathbf{z} \\
&= \int_{\mathbb{R}^n}\mathbb{V}(p_{\mathsf{M}}, p_{\mathsf{M}|\mathsf{Z}=\mathbf{z}})d\mu \\
&= \int_{\mathbb{R}^n}\mathbb{V}_{\mathsf{M}}(\mathbf{z})d\mu,
\end{aligned}$$

where $\mathbb{V}_{\mathsf{M}}(\mathbf{z}) = \mathbb{V}(p_{\mathsf{M}}, p_{\mathsf{M}|\mathsf{Z}=\mathbf{z}})$ and $d\mu = p_{\mathsf{Z}}(\mathbf{z})d\mathbf{z}$ is the probability measure associated to Z.

By using Lemma 2.7 in [22], we obtain

$$\mathbb{H}(\mathsf{M}) - \mathbb{H}(\mathsf{M}|\mathsf{Z}=\mathbf{z}) \leq \mathbb{V}_{\mathsf{M}}(\mathbf{z})\log\frac{|\mathcal{M}_n|}{\mathbb{V}_{\mathsf{M}}(\mathbf{z})}.$$

Multiplying by $p_{\mathsf{Z}}(\mathbf{z})$ and taking the integral, we find

$$\begin{aligned}
\mathbb{I}(\mathsf{M};\mathsf{Z}) &= \mathbb{H}(\mathsf{M}) - \mathbb{H}(\mathsf{M}|\mathsf{Z}) \\
&\leq \int_{\mathbb{R}^n}\mathbb{V}_{\mathsf{M}}(\mathbf{z})\log\frac{|\mathcal{M}_n|}{\mathbb{V}_{\mathsf{M}}(\mathbf{z})}d\mu \\
&= \int_{\mathbb{R}^n}\mathbb{V}_{\mathsf{M}}(\mathbf{z})\log|\mathcal{M}_n|d\mu - \int_{\mathbb{R}^n}\mathbb{V}_{\mathsf{M}}(\mathbf{z})\log\mathbb{V}_{\mathsf{M}}(\mathbf{z})d\mu.
\end{aligned}$$

From Jensen's inequality, using the fact that the function $t \mapsto t \log t$ is convex, we have that

$$\int_{\mathbb{R}^n} \mathbb{V}_{\mathsf{M}}(\mathbf{z}) \log \mathbb{V}_{\mathsf{M}}(\mathbf{z}) d\mu$$
$$\geq \left(\int_{\mathbb{R}^n} \mathbb{V}_{\mathsf{M}}(\mathbf{z}) d\mu \right) \log \left(\int_{\mathbb{R}^n} \mathbb{V}_{\mathsf{M}}(\mathbf{z}) d\mu \right)$$
$$= d_{\text{av}} \log d_{\text{av}}.$$

This completes the proof. □

Appendix II
Lattices that are good for coding

We recall some characterizations of "good" lattices for channel coding and shaping that have been proposed in the literature [15, 24, 28, 36]. A sequence of lattices is good for covering if its Voronoi region is asymptotically close to a sphere:

Definition 10 (Rogers-good). *Given a lattice Λ, let $r_{\text{cov}}(\Lambda)$ denote its covering radius and $r_{\text{eff}}(\Lambda)$ its effective radius (that is, the radius of a sphere having the same volume as the Voronoi region of Λ). A sequence of lattices $\Lambda^{(n)}$ is called* Rogers-good *or* covering-good *if $\lim_{n\to\infty} \frac{r_{\text{cov}}(\Lambda)}{r_{\text{eff}}(\Lambda)} = 1$.*

Definition 11 (Quantization-good). *A sequence of lattices $\Lambda^{(n)}$ is* quantization-good *if the normalized second moment $G(\Lambda^{(n)})$ tends to $\frac{1}{2\pi e}$ as n tends to infinity.*

Let us also introduce the notion of lattices which are good for the Gaussian channel without power constraint: [4]

Definition 12 (AWGN-good). *Given $\varepsilon > 0$ and an n-dimensional lattice Λ, let W^n be an i.i.d. Gaussian random vector of variance σ_ε^2 such that $\mathbb{P}\{\mathsf{W}^n \notin \mathcal{V}(\Lambda)\} = \varepsilon$. Consider the corresponding generalized SNR $\gamma_\Lambda(\sigma_\varepsilon) = \frac{(V(\Lambda))^{\frac{2}{n}}}{2\pi\sigma_\varepsilon^2}$. The sequence of lattices $\Lambda^{(n)}$ is* AWGN-good *if, for all $\varepsilon \in (0,1)$,*

$$\lim_{n\to\infty} \gamma_{\Lambda^{(n)}}(\sigma_\varepsilon) = e$$

and if, for a fixed generalized SNR greater than e, the quantity

$$\mathbb{P}\{\mathsf{W}^n \notin \mathcal{V}(\Lambda)\}$$

vanishes exponentially fast in n.

Observe that all that the previous properties are all invariant by scaling of the lattice.

Erez and Zamir [15] showed that lattice coding and decoding can achieve the capacity of the Gaussian channel. More precisely, one can prove the existence of a sequence of nested lattices $\Lambda_s^{(n)} \subset \Lambda_b^{(n)}$ such that

- the shaping lattice $\Lambda_s^{(n)}$ is simultaneously Rogers-good, quantization-good and AWGN-good,
- the fine lattice $\Lambda_b^{(n)}$ is AWGN-good.

When a random dither at the transmitter and an MMSE filter at the receiver are used, the Voronoi signal constellation $\Lambda_b^{(n)} \cap \mathcal{V}(\Lambda_s^{(n)})$ approaches the capacity of the mod-$\Lambda_s^{(n)}$ Gaussian channel, and consequently the capacity of the Gaussian channel, when n is large (see [15]).

Appendix III
Existence of good nested lattices: Proof of Proposition 4

Let $\mathcal{C}$ denote the set of $\mathbb{F}_p$-linear (n,k) codes, and let C be chosen uniformly at random from $\mathcal{C}$. Consider the corresponding Construction-A random lattice

$$\tilde{\Lambda}_s = \frac{1}{p} C + \mathbb{Z}^n.$$

By definition of the effective radius, we have:

$$p^k = \frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{\frac{n}{2}} r_{\text{eff}}(\tilde{\Lambda}_s)^n}.$$

We know from [34, Theorem 5] that with high probability, the lattice $\tilde{\Lambda}_s$ is Rogers, quantization and AWGN-good if the following properties are satisfied:

(i) $\exists \beta < \frac{1}{2} : \; k \leq \beta n$,
(ii) $\lim\limits_{n\to\infty} \frac{k}{\log^2 n} = \infty$,
(iii) $\forall n : \; r_{\min} < r_{\text{eff}}(\tilde{\Lambda}_s) < 2 r_{\min}$, where

$$r_{\min} = \min \left\{ \frac{1}{4}, \frac{(r_{\text{eff}}(\tilde{\Lambda}_s))^2}{32 n \sigma_b^2 E_P\left(\frac{r_{\text{eff}}(\tilde{\Lambda}_s)}{\sqrt{n}\sigma_b}\right)} \right\}.$$

In the previous formula, E_P denotes the Poltyrev exponent

$$E_P(\mu) = \begin{cases} \frac{1}{2}\left[(\mu-1) - \log \mu\right] & 1 < \mu \leq 2 \\ \frac{1}{2} \log \frac{e\mu}{4} & 2 \leq \mu \leq 4 \\ \frac{\mu}{8} & \mu \geq 4 \end{cases} \tag{38}$$

where $\mu = \frac{\gamma_{\Lambda_s}(\sigma_b)}{e}$. [5] Property (iii) implies that the fundamental volume is bounded by

$$\frac{\pi^{\frac{n}{2}} (r_{\min})^n}{\Gamma\left(\frac{n}{2}+1\right)} < V(\tilde{\Lambda}_s) = \frac{1}{p^k} < \frac{\pi^{\frac{n}{2}} (2 r_{\min})^n}{\Gamma\left(\frac{n}{2}+1\right)}, \tag{39}$$

which tends to 0 faster than exponentially, since Euler's Gamma function grows faster than any exponential. Given (n,k) with k satisfying (i) and (ii), consider $\tilde{p}(n,k)$ prime satisfying the condition (39). (The existence of such a prime number has been proven in [34].)

As explained in [34] (end of Section III), in order to use $\tilde{\Lambda}_s$ for power-constrained shaping it is necessary to scale it differently: we consider $\Lambda_s = ap\tilde{\Lambda}_s = \mathbf{B}_s \mathbb{Z}^n$ scaled so that its second moment satisfies $\sigma^2(\Lambda_s) = P$. As we have noted in Appendix II, such a scaling does not affect Rogers, quantization and AWGN-goodness.

Since Λ_s is quantization-good, its normalized second moment satisfies $G(\Lambda_s) = \frac{\sigma^2(\Lambda_s)}{V(\Lambda_s)^{\frac{2}{n}}} = \frac{P}{V(\Lambda_s)^{\frac{2}{n}}} \to \frac{1}{2\pi e}$ as $n \to \infty$. Therefore

$$V(\Lambda_s)^{\frac{2}{n}} = \frac{P}{G(\Lambda_s)} \to 2\pi P e.$$

For large n, we have

$$V(\Lambda_s) = a^n p^{n-k} \approx (2\pi e P)^{\frac{n}{2}}. \tag{40}$$

Since p^k grows superexponentially, so does p^{n-k} and we thus have $a \to 0$ and $ap \to \infty$ as $n \to \infty$. If we

[4] Our definition is the same as in [15, 37], except for the normalization factor 2π.

[5] Actually, the Poltyrev exponent is defined as a function of $\mu = V(\Lambda)^{\frac{2}{n}}/\sigma_b^2$ in [28] and as a function of $V(\Lambda)^{\frac{2}{n}}/(2\pi e \sigma_b^2)$ in [15].

set a in such a way that $V(\Lambda_s)$ is constant for $a \to 0$ and $p \to \infty$ (but may depend on n), then for each n we have a Minkowski-Hlawka type bound on the average behaviour of the theta series $\Theta_{\Lambda_s}(\tau)$ (see Lemma 3). Fix $\delta_n > 0$. For all n, there exists $\bar{p}(n,k,\delta_n,\tau)$ such that for every prime $p > \bar{p}(n,k,\delta_n,\tau)$ and the corresponding a,

$$\mathbb{E}\left[\Theta_{\Lambda_s}(\tau)\right] \leq 1 + \delta_n + \frac{1}{V(\Lambda_s)\tau^{\frac{n}{2}}}. \tag{41}$$

The following lemma, proven in Appendix IV, gives a more precise bound on the rate of convergence of the theta series to the Minkowski-Hlawka bound and guarantees that this choice of p is compatible with (39).

Lemma 9. *There exists a sequence $\delta_n \to 0$ such that for sufficiently large n, we have $\tilde{p}(n,k) > \bar{p}(n,k,\delta_n,y)$.*

Having defined the shaping lattice, we proceed with a nested code construction inspired by Section VII in [15]. Let C_b be chosen uniformly in the ensemble $\mathcal{C}_b$ of random linear (n,k_b) codes over $\mathbb{F}_q$, and denote by $\mathbf{A}_b$ its generator matrix. We know from [37] that if $\frac{n}{q} \to 0$, then the lattice

$$\Lambda_b = \mathbf{B}_s\left(\frac{1}{q}C_b + \mathbb{Z}^n\right).$$

is AWGN-good with high probability. Let $k_e < k_b$, and let $\mathbf{A}_e$ be the matrix whose columns are the first k_e columns of $\mathbf{A}_b$. This matrix generates an (n,k_e) linear code C_e over $\mathbb{F}_q$; note that averaging over the possible choices for C_b, this construction results in C_e being a uniformly chosen (n,k_e,q) linear code. We can consider the corresponding Construction-A lattice

$$\Lambda_e = \mathbf{B}_s\left(\frac{1}{q}C_e + \mathbb{Z}^n\right).$$

Clearly, we have $\Lambda_s \subseteq \Lambda_e \subseteq \Lambda_b$. As remarked in [37], there are many choices for q and k_e, k_b which ensure the properties

$$\begin{aligned} R'_n &= \frac{1}{n}\log\frac{V(\Lambda_s)}{V(\Lambda_e)} = \frac{k_e}{n}\log q \to R', \\ R_n &= \frac{1}{n}\log\frac{V(\Lambda_e)}{V(\Lambda_b)} = \frac{k_b}{n}\log q \to R. \end{aligned} \tag{42}$$

For example we can choose q to be the closest prime to $n\log n$ and define $k_e = \lfloor nR'(\log q)^{-1}\rfloor$, $k_b = \lfloor n(R+R')(\log q)^{-1}\rfloor$. Consider the expectation over over the sets $\mathcal{C}$ and $\mathcal{C}_e$ of (n,k,p) and (n,k_e,q) linear codes. By Lemma 1, we have:

$$\begin{aligned} &\lim_{n\to\infty}\mathbb{E}_{\mathcal{C},\mathcal{C}_e}\left[\epsilon_{\Lambda_e}(\sigma)\right] \\ &= \lim_{n\to\infty}\gamma_{\Lambda_e}(\sigma)^{\frac{n}{2}}\mathbb{E}_{\mathcal{C}}\left[\mathbb{E}_{\mathcal{C}_e}\left[\Theta_{\Lambda_e}\left(\frac{1}{2\pi\sigma^2}\right)\right]\right] - 1. \end{aligned} \tag{43}$$

Let $f(\mathbf{x}) = e^{-\pi\tau\|\mathbf{x}\|^2}$, $\bar{\mathbf{v}} = \mathbf{v} \bmod q$, and $C_e^* = C_e \setminus \{\mathbf{0}\}$. We have

$$\begin{aligned} &\mathbb{E}_{\mathcal{C}_e}\left[\Theta_{\Lambda_e}(\tau)\right] \\ &= \frac{1}{|\mathcal{C}_e|}\sum_{C_e\in\mathcal{C}_e}\left(\sum_{\substack{\mathbf{v}\in\mathbb{Z}^n\\ \bar{\mathbf{v}}=0}} f\left(\frac{\mathbf{B}_s\mathbf{v}}{q}\right) + \sum_{\substack{\mathbf{v}\in\mathbb{Z}^n\\ \bar{\mathbf{v}}\in C_e^*}} f\left(\frac{\mathbf{B}_s\mathbf{v}}{q}\right)\right) \\ &= \sum_{\mathbf{v}\in q\mathbb{Z}^n} f\left(\frac{\mathbf{B}_s\mathbf{v}}{q}\right) + \frac{q^{k_e}-1}{q^n-1}\sum_{\mathbf{v}\in\mathbb{Z}^n:\bar{\mathbf{v}}\neq 0} f\left(\frac{\mathbf{B}_s\mathbf{v}}{q}\right) \\ &= \sum_{\mathbf{v}\in\mathbb{Z}^n} f\left(\mathbf{B}_s\mathbf{v}\right) + \frac{q^{k_e}-1}{q^n-1}\sum_{\mathbf{v}\in\mathbb{Z}^n\setminus q\mathbb{Z}^n} f\left(\mathbf{B}_s\mathbf{v}/q\right) \\ &= \left(1 - \frac{q^{k_e}-1}{q^n-1}\right)\Theta_{\Lambda_s}(\tau) + \frac{q^{k_e}-1}{q^n-1}\Theta_{\Lambda_s}\left(\frac{\tau}{q^2}\right). \end{aligned}$$

In the last equation we have used the equality $\Theta_{a\Lambda}(\tau) = \Theta_\Lambda(a^2\tau)$.

We can now rewrite (43) as

$$\lim_{n\to\infty}\gamma_{\Lambda_e}(\sigma)^{\frac{n}{2}}\left(\mathbb{E}_{\mathcal{C}}\left[\Theta_{\Lambda_s}(\tau) + \frac{1}{q^{n-k_e}}\Theta_{\Lambda_s}\left(\frac{\tau}{q^2}\right)\right]\right) - 1$$

where $\tau = \frac{1}{2\pi\sigma^2}$. Using the property (41), this can be bounded by

$$\begin{aligned} &\lim_{n\to\infty}\gamma_{\Lambda_e}(\sigma)^{\frac{n}{2}}\left(1 + \frac{(2\pi\sigma^2)^{\frac{n}{2}}}{V(\Lambda_s)} + \delta_n\right) \\ &+ \lim_{n\to\infty}\gamma_{\Lambda_e}(\sigma)^{\frac{n}{2}}\left(\frac{1}{q^{n-k_e}}\left(1 + \frac{(2\pi\sigma^2q^2)^{\frac{n}{2}}}{V(\Lambda_s)} + \delta_n\right)\right) - 1 \\ &\leq \lim_{n\to\infty}\gamma_{\Lambda_e}(\sigma)^{\frac{n}{2}}\left(1 + \frac{1}{e^{nR'}\gamma_{\Lambda_e}(\sigma)^{\frac{n}{2}}} + \delta_n + \frac{1}{\gamma_{\Lambda_e}(\sigma)^{\frac{n}{2}}}\right) \\ &= \lim_{n\to\infty}\gamma_{\Lambda_e}(\sigma)^{\frac{n}{2}}(1+\delta_n) \end{aligned}$$

recalling that $e^{nR'_n} = q^{k_e}$ (see (42)). Therefore Λ_e is secrecy-good.

Further, we can show the majority of such lattices are secrecy-good. Fix $0 < c \leq \frac{1}{2}$ and let $\delta = \frac{\gamma_{\Lambda_e}(\sigma)^{\frac{n}{2}}(1+\delta_n)}{c}$. Then using Markov's inequality we get

$$\mathbb{P}\left\{\epsilon_{\Lambda_e}(\sigma) \geq \delta\right\} \leq \frac{\mathbb{E}\left[\epsilon_{\Lambda_e}(\sigma)\right]}{\delta} \leq c$$

Therefore if $\gamma_{\Lambda_e}(\sigma) < 1$, the sequence ${\Lambda_e}^{(n)}$ is secrecy-good with probability greater than $1 - c \geq \frac{1}{2}$.

To conclude, for n large enough there exists a set of measure going to 1 in the ensemble $\mathcal{C} \times \mathcal{C}_b$ such that Λ_s is Rogers, quantization and AWGN-good and Λ_b is AWGN-good [15], and a set of measure greater than $1/2$ in the same ensemble such that Λ_e is secrecy-good. The intersection of these sets being non-empty, the existence of a good sequence of nested lattices follows as stated.

Appendix IV
Proofs of technical lemmas

A. Proof of Lemma 3

Let $f(\mathbf{v}) = e^{-\pi\tau\|\mathbf{v}\|^2}$ for $\mathbf{v} \in \mathbb{R}^n$ and fixed $\tau \in \mathbb{R}^+$, and denote by C' the set of all nonzero codewords of C. Following [13], we have

$$\begin{aligned} &\frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\sum_{\mathbf{v}\in a\Lambda_C} f(\mathbf{v}) \\ &= \frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\left[\sum_{\mathbf{v}\in\mathbb{Z}^n:\bar{\mathbf{v}}=0} f(a\mathbf{v}) + \sum_{\mathbf{v}\in\mathbb{Z}^n:\bar{\mathbf{v}}\in C'} f(a\mathbf{v})\right] \\ &= \sum_{\mathbf{v}\in\mathbb{Z}^n:\bar{\mathbf{v}}=0} f(a\mathbf{v}) + \frac{p^k-1}{p^n-1}\sum_{\mathbf{v}\in\mathbb{Z}^n:\bar{\mathbf{v}}\neq 0} f(a\mathbf{v}) \end{aligned} \tag{44}$$

$$= \sum_{\mathbf{v}\in ap\mathbb{Z}^n} f(\mathbf{v}) + \frac{p^k-1}{p^n-1}\left(\sum_{\mathbf{v}\in a\mathbb{Z}^n} f(\mathbf{v}) - \sum_{\mathbf{v}\in ap\mathbb{Z}^n} f(\mathbf{v})\right) \tag{45}$$

where (44) is due to the balancedness of $\mathcal{C}$. We have

$$\sum_{\mathbf{v}\in ap\mathbb{Z}^n} f(\mathbf{v}) = \Theta_{ap\mathbb{Z}^n}(\tau) \to 1 \tag{46}$$

for any $\tau > 0$, since $ap \to \infty$ under the conditions given. Moreover,

$$\frac{p^k-1}{p^n-1}\sum_{\mathbf{v}\in a\mathbb{Z}^n} f(\mathbf{v}) \to V^{-1}\int_{\mathbb{R}^n} f(\mathbf{v})d\mathbf{v} \tag{47}$$

as $a \to 0$, $p \to \infty$ and $a^n p^{n-k} = V$ is fixed. To see this, consider any sequence $a_\ell \to 0$ and define $f_\ell(\mathbf{v}) = f\left(a_\ell \left\lfloor \frac{\mathbf{v}}{a_\ell}\right\rceil\right)$, then use Lebesgue's dominated convergence theorem, the functions f_ℓ being dominated by $g(\mathbf{v})$ which is equal to 1 if $\mathbf{v} \in \left[-\frac{1}{2}, \frac{1}{2}\right]^n$ and equal to $e^{-\pi\tau\sum_{i=1}^n \left(|v_i|-\frac{1}{2}\right)^2}$ otherwise. Thus, we have

$$\frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\sum_{\mathbf{v}\in a\Lambda_C} f(\mathbf{v}) \to 1 + V^{-1}\int_{\mathbb{R}^n} f(\mathbf{v})d\mathbf{v}. \tag{48}$$

Since $\int_{\mathbb{R}^n} f(\mathbf{v})d\mathbf{v} = \tau^{-n/2}$, we obtain (10).

Remark 13. Although we are primarily concerned with the theta series, the average behavior (48) is more general and may be of independent interest. In fact, (48) holds as long as the function $f(\cdot)$ satisfies conditions (46) and (47).

B. *Proof of the second part of Lemma 5*

Let $\varepsilon = \epsilon_{\Lambda'}(\sigma)$. From Lemma 4, we have that $\forall\bar{\lambda} \in \Lambda/\Lambda'$,

$$\frac{f_{\sigma,\bar{\lambda}}(\Lambda')}{f_{\sigma,\mathbf{0}}(\Lambda')} \in \left[\frac{1-\varepsilon}{1+\varepsilon}, 1\right].$$

Therefore, for all $\bar{\lambda} \in \Lambda/\Lambda'$:

$$\frac{|\Lambda/\Lambda'| \cdot f_{\sigma,\bar{\lambda}}(\Lambda')}{S} \in \left[\frac{1-\varepsilon}{1+\varepsilon}, \frac{1+\varepsilon}{1-\varepsilon}\right],$$

where $S = \sum_{\bar{\lambda}\in\Lambda/\Lambda'} f_{\sigma,\bar{\lambda}}(\Lambda')$. As a consequence:

$$\begin{aligned}
&|D_{\Lambda,\sigma}(\bar{\lambda}+\lambda') - p_{\mathsf{L}+\mathsf{L}'(\bar{\lambda}+\lambda')}| \\
&\quad = f_\sigma(\bar{\lambda}+\lambda')\left|\frac{1}{S} - \frac{1}{|\Lambda/\Lambda'|\, f_{\sigma,\bar{\lambda}}(\Lambda')}\right| \\
&\quad \le \frac{f_\sigma(\bar{\lambda}+\lambda')}{S}\max\left(\left|1-\frac{1+\varepsilon}{1-\varepsilon}\right|, \left|1-\frac{1-\varepsilon}{1+\varepsilon}\right|\right) \\
&\quad = \frac{2\varepsilon}{1-\varepsilon}D_{\Lambda,\sigma}(\bar{\lambda}+\lambda'). \qquad \square
\end{aligned}$$

C. *Proof of Lemma 9*

We study more explicitly the rate of convergence, by going back to the expression (45) in the proof of Lemma 3. We can rewrite it as

$$\begin{aligned}
&\left(1-\frac{p^k-1}{p^n-1}\right)\left(\Theta_{\mathbb{Z}^n}(a^2p^2\tau)\right) + \frac{p^k-1}{p^n-1}\left(\Theta_{\mathbb{Z}^n}(a^2\tau)\right) \\
&= \left(1-\frac{p^k-1}{p^n-1}\right)\left(\Theta_{\mathbb{Z}}(a^2p^2\tau)\right)^n + \frac{p^k-1}{p^n-1}\left(\Theta_{\mathbb{Z}}(a^2\tau)\right)^n
\end{aligned}$$

From the bound

$$\begin{aligned}
\int_{\mathbb{R}} e^{-\tau z^2}dz &= 2\int_0^\infty e^{-\tau z^2}dz \\
&\le \Theta_{\mathbb{Z}}(\tau) = 1 + 2\sum_{z\ge 1} e^{-yz^2} \\
&\le 1 + 2\int_0^\infty e^{-\tau z^2}dz = 1 + \int_{\mathbb{R}} e^{-\tau z^2}dz,
\end{aligned}$$

and recalling that $a^n p^{n-k} = V$, we find that

$$\begin{aligned}
\frac{1}{p^{n-k}}\left(\Theta_{\mathbb{Z}}(a^2\tau)\right)^n &\le \frac{a^n}{V}\left(1 + \frac{1}{a}\int_{\mathbb{R}} e^{-yz^2}dz\right)^n \\
&= \frac{1}{V}\int_{\mathbb{R}^n} e^{-\tau\|\mathbf{v}\|^2}d\mathbf{v} + o\left(\frac{1}{V^{1-\frac{1}{n}}p^{1-\frac{k}{n}}}\right),
\end{aligned}$$

while the lower bound is simply

$$\frac{1}{p^{n-k}}\left(\Theta_{\mathbb{Z}}(a\tau)\right)^n \ge \frac{1}{V}\int_{\mathbb{R}^n} e^{-\tau\|\mathbf{v}\|^2}d\mathbf{v}.$$

Similarly, we have

$$1 \le \left(\Theta_{\mathbb{Z}}(a^2p^2\tau)\right)^n \le 1 + \frac{V^{\frac{1}{n}}}{p^{\frac{k}{n}}}n\int_{\mathbb{R}} e^{-yz^2}dz + o\left(\frac{n^{\frac{1}{n}}}{p^{\frac{k}{n}}}\right).$$

It is not hard to see that the sequence $\tilde{q}(n,k)$ defined by (39) ensures (more than exponentially fast) convergence. □

References

[1] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, pp. 1355–1387, Oct. 1975.

[2] I. Csiszár, "Almost independence and secrecy capacity," *Problems of Information Transmission*, vol. 32, pp. 40–47, 1996.

[3] S. Goldwasser and S. Micali, "Probabilistic encryption," *J. Comput. Syst. Sci.*, vol. 28, no. 2, pp. 270–299, 1984.

[4] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel," in *Proc. CRYPTO 2012*, ser. Lecture Notes in Computer Science, vol. 7417. Springer-Verlag, pp. 294–311.

[5] A. Suresh, A. Subramanian, A. Thangaraj, M. Bloch, and S. McLaughlin, "Strong secrecy for erasure wiretap channels," in *IEEE Information Theory Workshop (ITW)*, 2010, pp. 1–5.

[6] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inform. Theory*, vol. 57, no. 10, pp. 6428–6443, Oct. 2011.

[7] D. Klinc, J. Ha, S. McLaughlin, J. Barros, and B.-J. Kwak, "LDPC codes for the Gaussian wiretap channel," *IEEE Trans. Information Forensics and Security*, vol. 6, pp. 532–540, Sept. 2011.

[8] L. Lai, H. El Gamal, and H. Poor, "The wiretap channel with feedback: Encryption over the channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5059–5067, nov. 2008.

[9] X. He and A. Yener, "Providing secrecy with lattice codes," in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, sept. 2008, pp. 1199–1206.

[10] F. Oggier, P. Solé, and J.-C. Belfiore, "Lattice codes for the wiretap Gaussian channel: Construction and analysis," Mar. 2011. [Online]. Available: http://arxiv.org/abs/1103.4086

[11] A. Ernvall-Hytonen and C. Hollanti, "On the eavesdropper's correct decision in Gaussian and fading wiretap channels using lattice codes," in *IEEE Information Theory Workshop (ITW)*, Oct. 2011, pp. 210–214.

[12] L.-C. Choo, C. Ling, and K.-K. Wong, “Achievable rates for lattice coding over the Gaussian wiretap channel,” in *ICC 2011 Physical Layer Security Workshop*, 2011.

[13] H. A. Loeliger, “Averaging bounds for lattices and linear codes,” *IEEE Trans. Inf. Theory*, vol. 43, pp. 1767–1773, Nov. 1997.

[14] G. Forney, M. Trott, and S.-Y. Chung, “Sphere-bound-achieving coset codes and multilevel coset codes,” *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 820–850, May 2000.

[15] U. Erez and R. Zamir, “Achieving 1/2 log(1+SNR) on the AWGN channel with lattice encoding and decoding,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, oct. 2004.

[16] F. R. Kschischang and S. Pasupathy, “Optimal non-uniform signaling for Gaussian channels,” *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 913–929, 1993.

[17] M. Bloch, “Achieving secrecy: Capacity vs. resolvability,” in *Proc. Int. Symp. Inform. Theory (ISIT 2011)*, St. Petersburg, Russia, July-August 2011.

[18] L. Luzzi and M. R. Bloch, “Capacity based random codes cannot achieve strong secrecy over symmetric wiretap channels,” in *SecureNets 2011*, 2011.

[19] D. Micciancio and O. Regev, “Lattice-based cryptography,” in *Post-quantum Cryprography*, D. J. Bernstein and J. Buchmann, Eds. Springer, 2008.

[20] M. Bloch and J. Barros, *Physical Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[21] Y. Liang, H. Poor, and S. Shamai, *Information Theoretic Security*. Foundations and Trends in Communications and Information Theory, Now Publishers, 2009.

[22] I. Csiszar and J. Korner, *Information Theory: coding theorems for discrete memoryless systems*. Akademiai Kiado, December 1981.

[23] P. M. Gruber and C. G. Lekkerkerker, *Geometry of Numbers*. Amsterdam, Netherlands: Elsevier, 1987.

[24] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices, and Groups*, 3rd ed. New York: Springer-Verlag, 1998.

[25] W. Banaszczyk, “New bounds in some transference theorems in the geometry of numbers,” *Math. Ann.*, vol. 296, pp. 625–635, 1993.

[26] D. Micciancio and O. Regev, “Worst-case to average-case reductions based on Gaussian measures,” in *Proc. Ann. Symp. Found. Computer Science*, Rome, Italy, Oct. 2004, pp. 372–381.

[27] J.-C. Belfiore, “Lattice codes for the compute-and-forward protocol: The flatness factor,” in *Proc. ITW 2011*, Paraty, Brazil, 2011.

[28] G. Poltyrev, “On coding without restrictions for the AWGN channel,” *IEEE Trans. Inf. Theory*, vol. 40, pp. 409–417, Mar. 1994.

[29] C. Gentry, C. Peikert, and V. Vaikuntanathan, “Trapdoors for hard lattices and new cryptographic constructions,” in *40th Annual ACM Symposium on Theory of Computing*, Victoria, Canada, 2008, pp. 197–206.

[30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[31] O. Regev, “On lattices, learning with errors, random linear codes, and cryptography,” *J. ACM*, vol. 56, no. 6, pp. 34:1–34:40, 2009.

[32] G. D. Forney, “On the role of MMSE estimation in approaching the information-theoretic limits of linear gaussian channels: Shannon meets Wiener,” in *Proceedings of the 41st Allerton Conference on Communication, Control and Computing*, 2003, pp. 430–439.

[33] M. R. Bloch and J. N. Laneman, “Secrecy from resolvability,” 2011. [Online]. Available: http://arxiv.org/abs/1105.5419

[34] U. Erez, S. Litsyn, and R. Zamir, “Lattices which are good for (almost) everything,” *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3401–3416, 2005.

[35] T. Han and S. Verdu, “Approximation theory of output statistics,” *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.

[36] R. Zamir and M. Feder, “On lattice quantization noise,” *IEEE Trans. Inf. Theory*, vol. 42, no. 4, pp. 1152–1159, 1996.

[37] B. Nazer and M. Gastpar, “Compute-and-forward: Harnessing interference through structured codes,” *IEEE Trans. Inf. Theory*, vol. 57, pp. 6463–6486, Oct. 2011.